



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 121370

To: Sarvamangala Devi
Location: REM 3C18
Art Unit: 1645
Thursday, May 06, 2004

Case Serial Number: 09/878781

From: Beverly Shears
Location: Remsen Bldg.
RM 1A54
Phone: 571-272-2528

beverly.shears@uspto.gov

Search Notes

Shears, Beverly

From: Devi, Sarvamangala
Sent: Wednesday, May 05, 2004 11:12 AM.
To: Shears, Beverly
Subject: 09/878,781

Good morning Beverly:

Please perform a sequence and an interference search for SEQ ID NO: 4 in application SN 09/878,781.

Thanks.

S. DEVI, Ph.D.
AU 1645
Rems - 3C18



STAFF USE ONLY

Date completed: 05-06-04
Searcher: Beverly c 2528
Terminal time: 23
Elapsed time: _____
CPU time: _____
Total time: 23
Number of Searches: _____
Number of Databases: 1

Search Site

____ STIC
____ CM-1
____ Pre-S

Type of Search

____ N.A. Sequence
____ A.A. Sequence
____ Structure
____ Bibliographic

Vendors

____ IG
____ STN
____ Dialog
____ APS
____ Geninfo
____ SDC
____ DARC/Questel
____ Other CGN

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 14:15:43 ; Search time 45 Seconds
(without alignments)
2355.869 Million cell updates/sec

Title: US-09-878-781-4

Perfect score: 1715
Sequence: 1 MVKVGVNGFGRIGRLAFRR.....EMSTYQVRLTYEPAKIAK 336

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP invertebrate:*
6: SP mammal:*
7: SP_mhc:*
8: SP_organelle:*
9: SP_phage:*
10: SP_plant:*
11: SP_rodent:*
12: SP_virus:*
13: SP_vertebrate:*
14: SP_unclassified:*
15: SP_virus:*
16: SP_bacteriap:*
17: SP_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1715	100.0	336	2	Q83ZF7 streptococ
2	1618	94.3	336	2	Q8XVU6 streptococ
3	1609	93.8	336	2	Q9ALM2 streptococ
4	1609	93.8	336	16	Q8E3B8 streptococ
5	1609	93.8	336	16	Q8DXS8 streptococ
6	1596	93.1	336	2	Q8GCR7 streptococ
7	1593.5	92.9	335	16	Q97NLI streptococ
8	1593.5	92.9	335	16	Q8CWN6 streptococ
9	1584	92.4	336	2	Q8KHG1 streptococ
10	1579.5	92.1	337	16	Q8DVT3 streptococ
11	1578	92.0	336	2	Q8XVU5 streptococ
12	1572	91.7	336	2	Q8VVB9 streptococ
13	1518	88.5	320	2	Q9L5X6 streptococ
14	1468.5	85.6	309	2	Q9AUT7 streptococ
15	1467.5	85.6	309	2	Q9AUT9 streptococ
16	1465.5	85.5	309	2	Q9AUT4 streptococ

17	1463.5	85.3	309	2	Q9AUT5 streptococ
18	1462.5	85.3	308	2	Q9AUT8 streptococ
19	1445.5	84.3	305	2	Q9AUT6 streptococ
20	1410	82.2	336	16	Q9CDH4 streptococ
21	1337.5	78.0	333	16	Q83318 streptococ
22	1262	73.6	336	16	Q8V411 streptococ
23	1261	73.5	336	16	Q928H9 streptococ
24	1227.5	71.6	332	16	Q8XKT9 streptococ
25	1218.5	71.0	335	16	Q8RPN9 streptococ
26	1197.5	69.8	334	16	Q9AUT8 streptococ
27	1190	69.4	336	16	Q8CPY5 streptococ
28	1189.5	69.4	330	2	Q83XK3 streptococ
29	1186.5	69.2	330	2	Q84H26 streptococ
30	1185.5	69.1	330	2	Q84H24 streptococ
31	1184.5	69.1	334	16	Q9UX95 streptococ
32	1183	69.0	336	16	Q925C5 streptococ
33	1182.5	69.0	330	2	Q84H25 streptococ
34	1178.5	68.7	335	2	Q93M61 streptococ
35	1138	66.4	311	2	Q8L2R0 streptococ
36	1132.5	66.0	310	2	Q8L2P7 streptococ
37	1124.5	65.6	310	2	Q8L2P6 streptococ
38	1124	65.5	311	2	Q8L2P8 streptococ
39	1123.5	65.5	310	2	Q8L2O6 streptococ
40	1121.5	65.4	311	2	Q8VW79 streptococ
41	1120.5	65.3	310	2	Q8L2O3 streptococ
42	1120.5	65.3	310	2	Q8L2P5 streptococ
43	1118.5	65.2	310	2	Q8L2Q2 streptococ
44	1116.5	65.1	310	2	Q8L2R7 streptococ
45	1115.5	65.0	310	2	Q8L2R2 streptococ

ALIGNMENTS

RESULT 1

ID	Q83ZF7	PRELIMINARY:	PRT:	336 AA.
AC	Q83ZF7			
DT	01-JUN-2003 (TREMBLrel. 24, Created)			
DT	01-JUN-2003 (TREMBLrel. 24, Last sequence update)			
DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)			
DE	Glyceraldehyde 3-P dehydrogenase.			
GN	GAPC.			
OS	Streptococcus dysgalactiae.			
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;			
OC	Streptococcus.			
OX	NCBI_TaxID=1334;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Peres-Casal J.F., Bolton A., Song X.-M., Willson P., Potter A.A.;			
RT	"Use of the surface proteins GapC and Mig of Streptococcus			
RT	dysgalactiae as protective antigens against mastitis in non-lactating			
RT	cows."			
RU	Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF375662; AAP31408.1; -			
DR	GO; GO:0004365; P:glyceraldehyde-3-phosphate dehydrogenase (p...; IEA.			
DR	GO; GO:0006096; P:glycolysis; IEA.			
DR	InterPro: IPR006424; GAPDH-1.			
DR	InterPro: IPR000173; GAP_dhhydrogenase.			
DR	Pfam: PF00044; gpdh. 1.			
DR	Pfam: PF02800; gpdh. C. 1.			
DR	PRINTS; PR00078; G3PDHCRNAS.			
DR	TIGRFAMs; TIGR01534; GAPDH-I; 1.			
DR	PROSITE; PS00071; GAPDH; 1.			
DR	SEQUENCE 336 AA; 35928 MW; 11828218CF037076 CRC64;			

Query Match 100.0%; Score 1715; DB 2; Length 336;

Best Local Similarity 100.0%; Pred. No. 5 7e-107; Indels 0; Gaps 0;

Matches	336;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	1	MVKVGVNGFGRIGRLAFRRIONVGVETVRINDLIDPVMALHLIKYDTQGRFDIVAV	60						
DB	1	MVKVGVNGFGRIGRLAFRRIONVGVETVRINDLIDPVMALHLIKYDTQGRFDIVAV	60						

QY 61 KEGGEVNGNFIKVSARDEPNDIATDGVETIYATGPAKKAEMKHJHANGAKKVI 120
 DB 61 KEGGEVNGNFIKVSARDEPNDIATDGVETIYATGPAKKAEMKHJHANGAKKVI 120
 QY 121 TAGGNDVKTIVNTHNDILDTGTVISGASCTTNCLAPMAKALDPAFGIOGKMTTHA 180
 DB 121 TAGGNDVKTIVNTHNDILDTGTVISGASCTTNCLAPMAKALDPAFGIOGKMTTHA 180
 QY 181 YTGDMQMLDGPFRGGDLRRARAGANIVPNSGTAAKAIIGVIPLENGKLDGAAGRPVPT 240
 DB 181 YTGDMQMLDGPFRGGDLRRARAGANIVPNSGTAAKAIIGVIPLENGKLDGAAGRPVPT 240
 QY 241 GSVTELVVTLIDKXVSVDENINAMKAASNDSPGTEPDIYSSDIYVSGSLFPDATTQTKM 300
 DB 241 GSVTELVVTLIDKXVSVDENINAMKAASNDSPGTEPDIYSSDIYVSGSLFPDATTQTKM 300
 QY 301 EVDGSQLVXVSVNEMSYTAOLVTLTEYFAKIAK 336
 DB 301 EVDGSQLVXVSVNEMSYTAOLVTLTEYFAKIAK 336

RESULT 2

Q8KVU6 PRELIMINARY; PRT; 336 AA.

ID Q8KVU6
 AC Q8KVU6;
 DT 01-OCT-2002 (TREMBLrel. 22, Created)
 DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Glyceraldehyde-3-phosphate dehydrogenase.
 OS Streptococcus uberis.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 CC Streptococcus.
 CX NCBI_TaxID=1349;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=ATCC 9927;
 RA Fontaine M.C., Perez-Casal J., Song X.-M., Sheldford J., Willison P.J.,
 RA Potter A.A.;
 RT "Immune reaction of dairy cattle with recombinant GapC and chimeric CAMP
 RT antigens confers protection against heterologous challenge with
 RT Streptococcus uberis."; the EMBL/GenBank/DBJ databases.
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
 CC DEHYDROGENASE FAMILY.
 DR EMBL: AF421390; AA013711.1;
 DR GO: GO:000435; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .); IEA.
 DR GO: GO:0016491; F:oxidoreductase activity; IEA.
 DR GO: GO:0006096; P:glycolysis; IEA.
 DR InterPro: IPR006424; GAPDH-I.
 DR InterPro: IPR001173; GAP_dhhydrogenase.
 DR Pfam: PF00044; spdh. 1.
 DR Pfam: PF02800; spdh. 1.
 DR PRINTS: PR00078; G3BDHGNAS.
 DR TIGRFAMs: TIGR01534; GAPDH-I.1.
 DR PROSITE: PS00071; GAPDH; 1.
 KW Oxidoreductase.
 SQ SEQUENCE 336 AA; 35906 MW; 755D74A4548B61D4 CRC64;

Query Match 94.3%; Score 1618; DB 2; Length 336;
 Best local similarity 92.9%; Pred. No. 1.8e-100;
 Matches 312; Conservative 15; Mismatches 9; Indels 0; Gaps 0;

QY 1 MYVKGINGFGRLGLAFRIQNVGEVTRINDLTPNMLAHLYKTYTQGRPGTIVY 60
 DB 1 MYVKGINGFGRLGLAFRIQNVGEVTRINDLTPNMLAHLYKTYTQGRPGTIVY 60
 QY 61 KEGGEVNGNFIKVSARDEPNDIATDGVETIYATGPAKKAEMKHJHANGAKKVI 120
 DB 61 KEGGEVNGNFIKVSARDEPNDIATDGVETIYATGPAKKAEMKHJHANGAKKVI 120
 QY 121 TAGGNDVKTIVNTHNDILDTGTVISGASCTTNCLAPMAKALDPAFGIOGKMTTHA 180
 DB 121 TAGGNDVKTIVNTHNDILDTGTVISGASCTTNCLAPMAKALDPAFGIOGKMTTHA 180

DB 121 TAGGNDVKTIVNTHNDILDTGTVISGASCTTNCLAPMAKALDPAFGIOGKMTTHA 180
 QY 181 YTGDMQMLDGPFRGGDLRRARAGANIVPNSGTAAKAIIGVIPLENGKLDGAAGRPVPT 240
 DB 181 YTGDMQMLDGPFRGGDLRRARAGANIVPNSGTAAKAIIGVIPLENGKLDGAAGRPVPT 240
 QY 241 GSVTELVVTLIDKXVSVDENINAMKAASNDSPGTEPDIYSSDIYVSGSLFPDATTQTKM 300
 DB 241 GSVTELVVTLIDKXVSVDENINAMKAASNDSPGTEPDIYSSDIYVSGSLFPDATTQTKM 300
 QY 301 EVDGSQLVXVSVNEMSYTAOLVTLTEYFAKIAK 336
 DB 301 EVDGSQLVXVSVNEMSYTAOLVTLTEYFAKIAK 336

RESULT 3

Q9ALM2 PRELIMINARY; PRT; 336 AA.

ID Q9ALM2
 AC Q9ALM2;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Glyceraldehyde-3-phosphate dehydrogenase.
 OS Streptococcus agalactiae.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 CC Streptococcus.
 CX NCBI_TaxID=1311;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=J48;
 RA Selfert K.N., Blaisieis A.S., McArthur W.P., Brady L.J.;
 RA "The Group B Streptococcal Surface Antigen Delta is a Glyceraldehyde
 RT 3-Phosphate Dehydrogenase."; the EMBL/GenBank/DBJ databases.
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
 CC DEHYDROGENASE FAMILY.
 DR EMBL: AF338416; AA014387.1;
 DR HSP; P00362; IGD1.
 DR GO: GO:000435; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .); IEA.
 DR GO: GO:0016491; F:oxidoreductase activity; IEA.
 DR GO: GO:0006096; P:glycolysis; IEA.
 DR InterPro: IPR006424; GAPDH-I.
 DR InterPro: IPR001173; GAP_dhhydrogenase.
 DR Pfam: PF00044; spdh. 1.
 DR Pfam: PF02800; spdh. 1.
 DR PRINTS: PR00078; G3BDHGNAS.
 DR TIGRFAMs: TIGR01534; GAPDH-I.1.
 DR PROSITE: PS00071; GAPDH; 1.
 KW Oxidoreductase.
 SQ SEQUENCE 336 AA; 36005 MW; 982E8M05CA343C9 CRC64;

Query Match 93.8%; Score 1609; DB 2; Length 336;
 Best local similarity 92.0%; Pred. No. 7.2e-100;
 Matches 309; Conservative 17; Mismatches 10; Indels 0; Gaps 0;

QY 1 MYVKGINGFGRLGLAFRIQNVGEVTRINDLTPNMLAHLYKTYTQGRPGTIVY 60
 DB 1 MYVKGINGFGRLGLAFRIQNVGEVTRINDLTPNMLAHLYKTYTQGRPGTIVY 60
 QY 61 KEGGEVNGNFIKVSARDEPNDIATDGVETIYATGPAKKAEMKHJHANGAKKVI 120
 DB 61 KEGGEVNGNFIKVSARDEPNDIATDGVETIYATGPAKKAEMKHJHANGAKKVI 120
 QY 121 TAGGNDVKTIVNTHNDILDTGTVISGASCTTNCLAPMAKALDPAFGIOGKMTTHA 180
 DB 121 TAGGNDVKTIVNTHNDILDTGTVISGASCTTNCLAPMAKALDPAFGIOGKMTTHA 180
 QY 181 YTGDMQMLDGPFRGGDLRRARAGANIVPNSGTAAKAIIGVIPLENGKLDGAAGRPVPT 240
 DB 181 YTGDMQMLDGPFRGGDLRRARAGANIVPNSGTAAKAIIGVIPLENGKLDGAAGRPVPT 240
 QY 241 GSVTELVVTLIDKXVSVDENINAMKAASNDSPGTEPDIYSSDIYVSGSLFPDATTQTKM 300
 DB 241 GSVTELVVTLIDKXVSVDENINAMKAASNDSPGTEPDIYSSDIYVSGSLFPDATTQTKM 300

Db 241 GSVTELVATLEKDVTEVEVNAMKAANDSYGTEDPVSSDIWJISYGLFDPATQTKVQ 300
 QY 301 EVDGSQLVYVSWYNDENMSYTAQVLTLEYFAKIAK 336
 Db 302 TVDGNQLVKVVSWYNDENMSYTAQVLTLEYFAKIAK 336

RESULT 4

Q8E3E8 PRELIMINARY; PRT: 336 AA.

ID Q8E3E8 01-MAR-2003 (TREMblrel. 23, Created)
 AC Q8E3E8 01-MAR-2003 (TREMblrel. 23, Last sequence update)
 DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
 DE Glyceraldehyde 3-phosphate dehydrogenase.
 GN G8S1811.
 OS Streptococcus agalactiae (serotype III).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 CX NCBI_TaxID=216495;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NE9316 / Serotype III;
 RX MEDLINE=22242508; PubMed=12354221;
 RA Glaeser P., Ruesink C., Buchliesser C., Chevalier F., Frangoul L.,
 RA Masdek T., Zouine M., Couve E., Talhouk L., Poyart C., Trieu-Chot P.,
 RA Kunst F.;
 RT "Genome sequence of Streptococcus agalactiae, a pathogen causing
 RT invasive neonatal disease."
 RL Mol. Microbiol. 45:1499-1511 (2002).
 DR EMBL; AL765853; CAD7470.1; -.
 DR GO; GO:0053365; F:glyceraldehyde-3-phosphate dehydrogenase (p. .; IEA.
 DR SAGALIST; gbs1811; -.
 DE GO; GO:0006396; P:glycolysis; IEA.
 DR Interpro; IPR006424; GAPDH-I.
 DR Interpro; IPR001173; GAP_dhrogenase.
 DR Pfam; PF02800; gpdh_C.1.
 DR PRINTS; PR00078; G3PDHDSGNASE.
 DR TIGRPFAM; TIGR01534; GAPDH-I; 1.
 DR PROSITE; PS00071; GAPDH; 1.
 KM Complete proteome.
 SQ SEQUENCE 336 AA; 36005 MW; 982E18A05CA343C9 CRC64;

Query Match 93.8%; Score 1609; DB 16; Length 336;
 Best Local Similarity 92.0%; Pred. No. 7.2e-100;
 Matches 309; Conservative 17; Mismatches 10; Indels 0; Gaps 0;

QY 1 MVTYVINGEGRIGR-LAFRIQNVGEVTRINDLTPMMLAHLLKXDTTQGRFGTVEV 60
 Db 1 MVTYVINGEGRIGR-LAFRIQNVGEVTRINDLTPMMLAHLLKXDTTQGRFGTVEV 60
 QY 61 KEGFEVNGAFIKVSAERDPENIDMTGVEIVLEATGFAKKAERKHLHANGAKKVI 120
 Db 61 KEGFEVNGAFIKVSAERDPENIDMTGVEIVLEATGFAKKAERKHLHANGAKKVI 120
 QY 121 TAPGANDVKTVPNNHNDLIDGTEVYISGASCTTNCIAPMAKALHDPAGIOMGLMTTIIA 180
 Db 121 TAPGANDVKTVPNNHNDLIDGTEVYISGASCTTNCIAPMAKALHDPAGIOMGLMTTIIA 180
 QY 181 YTGDMILLDGPFRGGDLRRARAGANIVPNSGAAGAIGLVIPELNGKLDGAAGRPVPT 240
 Db 181 YTGDMILLDGPFRGGDLRRARAGANIVPNSGAAGAIGLVIPELNGKLDGAAGRPVPT 240
 QY 241 GSVTELVATLEKDVTEVEVNAMKAANDSYGTEDPVSSDIWJISYGLFDPATQTKVQ 300
 Db 241 GSVTELVATLEKDVTEVEVNAMKAANDSYGTEDPVSSDIWJISYGLFDPATQTKVQ 300
 QY 301 EVDGSQLVYVSWYNDENMSYTAQVLTLEYFAKIAK 336
 Db 301 TVDGNQLVKVVSWYNDENMSYTAQVLTLEYFAKIAK 336

RESULT 5

Q8DXS8 PRELIMINARY; PRT: 336 AA.

ID Q8DXS8 01-MAR-2003 (TREMblrel. 23, Created)
 AC Q8DXS8 01-MAR-2003 (TREMblrel. 23, Last sequence update)
 DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
 DE Glyceraldehyde 3-phosphate dehydrogenase.
 GN GAP OR SAG1768.
 OS Streptococcus agalactiae (serotype V).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 CX NCBI_TaxID=216466;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=2603 V/R / Serotype V;
 RX MEDLINE=22222988; PubMed=12200547;
 RA Terebin H., Masiqant V., Cieslewicz M.J., Elen J.A., Peterson S.,
 RA Messels M.R., Paulsen I.T., Nelson K.E., Margalit I., Read T.D.,
 RA Madoff L.C., Wolf A.M., Beaman M.J., Brinkac R., Daugherty S.C.,
 RA Deboy R.T., Durkin A.S., Kolonay J.F., Madupu R., Lewis M.R.,
 RA Radune D., Fedorova N.B., Scanlan D., Kouri H., Mulligan S.,
 RA Carty H.A., Cline R.T., Van Aken S.E., Gill U., Scarcelli M., Mora M.,
 RA Jacobini E.T., Breckon C., Galli G., Mariani M., Vegen F., Malone D.,
 RA Rinaldo D., Rappuoli R., Telford J.L., Kasper D.L., Grandi G.,
 RA Fraser C.M.;
 RT "Complete genome sequence and comparative genomic analysis of an
 RT emerging human pathogen, serotype V Streptococcus agalactiae."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396 (2002).
 DR EMBL; AE014272; AAN00631.1; -.
 DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. .; IEA.
 DR GO; GO:0006096; P:glycolysis; IEA.
 DR Interpro; IPR006424; GAPDH-I.
 DR Interpro; IPR001173; GAP_dhrogenase.
 DR Pfam; PF00044; gpdh_C.1.
 DR PRINTS; PR00078; G3PDHDSGNASE.
 DR TIGRPFAM; TIGR01534; GAPDH-I; 1.
 DR PROSITE; PS00071; GAPDH; 1.
 KM Complete proteome.
 SQ SEQUENCE 336 AA; 36005 MW; 982E18A05CA343C9 CRC64;

Query Match 93.8%; Score 1609; DB 16; Length 336;
 Best Local Similarity 92.0%; Pred. No. 7.2e-100;
 Matches 309; Conservative 17; Mismatches 10; Indels 0; Gaps 0;

QY 1 MVTYVINGEGRIGR-LAFRIQNVGEVTRINDLTPMMLAHLLKXDTTQGRFGTVEV 60
 Db 1 MVTYVINGEGRIGR-LAFRIQNVGEVTRINDLTPMMLAHLLKXDTTQGRFGTVEV 60
 QY 61 KEGFEVNGAFIKVSAERDPENIDMTGVEIVLEATGFAKKAERKHLHANGAKKVI 120
 Db 61 KEGFEVNGAFIKVSAERDPENIDMTGVEIVLEATGFAKKAERKHLHANGAKKVI 120
 QY 121 TAPGANDVKTVPNNHNDLIDGTEVYISGASCTTNCIAPMAKALHDPAGIOMGLMTTIIA 180
 Db 121 TAPGANDVKTVPNNHNDLIDGTEVYISGASCTTNCIAPMAKALHDPAGIOMGLMTTIIA 180
 QY 181 YTGDMILLDGPFRGGDLRRARAGANIVPNSGAAGAIGLVIPELNGKLDGAAGRPVPT 240
 Db 181 YTGDMILLDGPFRGGDLRRARAGANIVPNSGAAGAIGLVIPELNGKLDGAAGRPVPT 240
 QY 241 GSVTELVATLEKDVTEVEVNAMKAANDSYGTEDPVSSDIWJISYGLFDPATQTKVQ 300
 Db 241 GSVTELVATLEKDVTEVEVNAMKAANDSYGTEDPVSSDIWJISYGLFDPATQTKVQ 300
 QY 301 EVDGSQLVYVSWYNDENMSYTAQVLTLEYFAKIAK 336
 Db 301 TVDGNQLVKVVSWYNDENMSYTAQVLTLEYFAKIAK 336

RESULT 6

08GCR7
ID 08GCR7 PRELIMINARY; PRT; 336 AA.
AC 08GCR7;
DT 01-VAR-2003 (TREMELrel. 23, Created)
DT 01-VAR-2003 (TREMELrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12).
OS Streptococcus suis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1307;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=8735;
RA Brassard V., Gottschalk M., Quessey S.;
RT "Cloning and purification of Streptococcus suis serotype 2
RT glyceraldehyde-3-phosphate dehydrogenase.";
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY167026; AAB6058.1; -
DR GO: GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.
DR GO: GO:0016491; F:oxidoreductase activity; IEA.
DR GO: GO:0006096; P:glycolysis; IEA.
DR InterPro: IPR006424; GAPDH-I.
DR InterPro: IPR000173; GAP_dhydrogenase.
DR Pfam: PF00044; gpdh; 1.
DR Pfam: PF02800; gpdh; 1.
DR PRINTS: PR00078; G3PDHGRNASF.
DR TIGRFAMs: TIGR01534; GAPDH-I; 1.
DR PROSITE: PS00071; GAPDH; 1.
KW Oxidoreductase.
SQ SEQUENCE 336 AA; 3582 MW; 1DBB9B1A92DF59 CRC64;

Query Match 93.1%; Score 1596; DB 2; Length 336;
Best Local Similarity 91.1%; Pred. No. 5,4e-99;
Matches 306; Conservative 17; Mismatches 13; Indels 0; Gaps 0;

QY 1 MVMKVGINGFGRIGRLAFRIQNVGEVTRINDLTDPMALHLKDTTQGRDGTVEV 60
DB 1 MVMKVGINGFGRIGRLAFRIQNVGEVTRINDLTDPMALHLKDTTQGRDGTVEV 60
QY 61 KEGFEVNGKFKYSARDEPNDIMATDGVLEATGFPKKAARKHLANGAKKVI 120
DB 61 KEGFEVNGKFKYSARDEPNDIMATDGVLEATGFPKKAARKHLANGAKKVI 120
QY 121 TARGNDVKTIVENTNHDLDGTETVSGASCTTNCLAPAKALHDHFGQKMTTTHA 180
DB 121 TARGNDVKTIVENTNHDLDGTETVSGASCTTNCLAPAKALHDHFGQKMTTTHA 180
QY 181 YTGDMILDGPHRGDLRRAPAGANIVPNSGAAGAIGVYIPELNGKLDGAARVPPT 240
DB 181 YTGDMILDGPHRGDLRRAPAGANIVPNSGAAGAIGVYIPELNGKLDGAARVPPT 240
QY 241 GSVELVLTLDKNSVDEINAAKASNDSEFTEDPIYSSDVGVSYSGLFDATOTKM 300
DB 241 GSVELVLTLDKNSVDEINAAKASNDSEFTEDPIYSSDVGVSYSGLFDATOTKM 300
QY 301 EVDGSQLVKKVSWYNENSYTAQVLTLETFPAKIAK 336
DB 301 EVDGSQLVKKVSWYNENSYTAQVLTLETFPAKIAK 336

RESULT 7
097NLI PRELIMINARY; PRT; 335 AA.
AC 097NLI;
DT 01-OCT-2001 (TREMELrel. 18, Created)
DT 01-OCT-2001 (TREMELrel. 18, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12).
GN SP2012 OR GAPDH.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.

OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC BAA-334 / TIGR4;
RX MEDLINE=2157209; PubMed=11463916;
RA Tettelin H., Nelson K.B., Paulsen I.T., Eisen J.A., Read T.D.,
RA Peterson S., Heidelberg J., Deboy R.T., Haft D.H., Dodson R.J.,
RA Durkin A.S., Gwin M., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA Holtzapple B., Knouri H., Wolf A.M., Uterback T.R., Hansen C.L.,
RA McDonald L.A., Feldblyum T.V., Anguilo S., Dickenson T., Hickey E.K.,
RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.M.,
RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT "Complete genome sequence of a virulent isolate of Streptococcus
RT pneumoniae";
RL Science 293:498-506(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC1733;
RA Bergmann S., Hammerschmidt S.;
RT "Identification of pneumococcal GAPDH as plasmin(ogen)-binding
RT protein." (AUG-2002) to the EMBL/GenBank/DBJ databases.
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AE007490; AK76079.1; -
DR EMBL: AJ505822; CAD44376.1; -
DR PIR: F95235; P95235.
DR HSP: P00354; 3GPD.
DR TIGR: SP0012; -
DR GO: GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.
DR GO: GO:0006096; P:glycolysis; IEA.
DR InterPro: IPR006424; GAPDH-I.
DR InterPro: IPR000173; GAP_dhydrogenase.
DR Pfam: PF00044; gpdh; 1.
DR Pfam: PF02800; gpdh; 1.
DR PRINTS: PR00078; G3PDHGRNASF.
DR TIGRFAMs: TIGR01534; GAPDH-I; 1.
DR PROSITE: PS00071; GAPDH; 1.
KW Oxidoreductase; Complete proteome.
SQ SEQUENCE 335 AA; 3585 MW; DA483CEA423E747B CRC64;

Query Match 92.9%; Score 1593.5; DB 16; Length 335;
Best Local Similarity 92.6%; Pred. No. 7.9e-99;
Matches 311; Conservative 12; Mismatches 12; Indels 1; Gaps 1;

QY 1 MVMKVGINGFGRIGRLAFRIQNVGEVTRINDLTDPMALHLKDTTQGRDGTVEV 60
DB 1 MVMKVGINGFGRIGRLAFRIQNVGEVTRINDLTDPMALHLKDTTQGRDGTVEV 60
QY 61 KEGFEVNGKFKYSARDEPNDIMATDGVLEATGFPKKAARKHLANGAKKVI 120
DB 61 KEGFEVNGKFKYSARDEPNDIMATDGVLEATGFPKKAARKHLANGAKKVI 120
QY 121 TARGNDVKTIVENTNHDLDGTETVSGASCTTNCLAPAKALHDHFGQKMTTTHA 180
DB 121 TARGNDVKTIVENTNHDLDGTETVSGASCTTNCLAPAKALHDHFGQKMTTTHA 180
QY 181 YTGDMILDGPHRGDLRRAPAGANIVPNSGAAGAIGVYIPELNGKLDGAARVPPT 240
DB 181 YTGDMILDGPHRGDLRRAPAGANIVPNSGAAGAIGVYIPELNGKLDGAARVPPT 240
QY 241 GSVELVLTLDKNSVDEINAAKASNDSEFTEDPIYSSDVGVSYSGLFDATOTKM 300
DB 241 GSVELVLTLDKNSVDEINAAKASNDSEFTEDPIYSSDVGVSYSGLFDATOTKM 300
QY 301 EVDGSQLVKKVSWYNENSYTAQVLTLETFPAKIAK 336
DB 301 EVDGSQLVKKVSWYNENSYTAQVLTLETFPAKIAK 336

RESULT 8
08CWN6 PRELIMINARY; PRT; 359 AA.
AC 08CWN6;
GN SP2012 OR GAPDH.

DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Glyceraldehyde 3-phosphate dehydrogenase [phosphorylating]
 DE (EC 1.2.1.12)
 GN GAPD OR SPB1825.
 OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).
 OC Bacteria; Firmicutes; Iactobacillales; Streptococcaceae;
 OC Streptococcus.
 NCBI_TaxID=171101;
 RN
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21429245; PubMed=11544234;
 RA Hoskins J., Albom W.E. Jr., Arnold J., Blaszcak L.C., Burgess S.,
 RA Dehoff B.S., Estrem S.T., Fritz L., Fu D.-T., Fuller W., Geringer C.,
 RA Gilmore R., Glass J.S., Klotz H., Kraft A.R., Lagace R.E.,
 RA Liebman D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P.,
 RA McHaren S.M., McHenry M., Mcleaster K., Mundy C.W., Nicas T.I.,
 RA Norris F.H., O'Gara M., Peery R.B., Robertson G.T., Rocky P.,
 RA Sun P.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,
 RA Zook C.A., Baltz R.H., Taskunas S.R., Rostock P.R. Jr., Skarnd P.L.,
 RA Glas J.I.,
 RT "Genome of the bacterium Streptococcus pneumoniae strain R6."
 RL J. Bacteriol. 183:5709-5717(2001).
 DR EMBL: A308547; AAL0628.1; -.
 DR PIR: G98099; G98099.
 DR GO: GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.
 DR GO: GO:0016491; F:oxidoreductase activity; IEA.
 DR GO: GO:0006096; P:glycolysis; IEA.
 DR InterPro: IPR006424; GAPDH-1.
 DR InterPro: IPR00173; GAP_dhrogenase.
 DR Pfam: PF02800; gpdh. C. 1.
 DR PRINTS: PR00076; G3PDHGNASE.
 DR TIGRFAMs: TIGR01534; GAPDH-1; 1.
 DR PROSITE: PS00071; GAPDH. 1.
 DR Oxidoreductase; Complete proteome.
 SQ SEQUENCE 359 AA; 38763 MW; 61EF0E35B30B60 CRC64;

Query Match 92.9%; Score 1593.5; DB 16; Length 359;
 Best Local Similarity 92.6%; Pred. No. 8.7e-99;
 Matches 311; Conservative 12; Mismatches 12; Indels 1; Gaps 1;

QY 1 M V V K G I N G F R I G R I A F R I O N V G E V T R I N D L T D P M L A H L K Y D T T G R F D G T V E V 60
 DB 25 M V V K V I N G F R I G R I A F R I O N V G E V T R I N D L T D P M L A H L K Y D T T G R F D G T V E V 84
 QY 61 K E G F E V N G N F I K V S A R D E N I M A T D G V E I V L E A T G F P A K K A A E K H L H A N G A K K V I 120
 DB 85 K E G F E V N G K F I K V S A R D E P Q I D M A T D G V E I V L E A T G F P A K K A A E K H L K G A K K V I 143
 QY 121 T A P G S D V K V V N T N H D I L D G T E V I S G A S C T T N C L A P M A K A L D A P G I O K G I M T T H A 180
 DB 144 T A P G S D V K V V N T N H D I L D G T E V I S G A S C T T N C L A P M A K A L D A P G I O K G I M T T H A 203
 QY 181 Y T G D Q M I L D G P H R G D L R R A R A G A N I V P N S T G A A K A I G L V I P E I N K L D G A A Q R V P V P T 240
 DB 204 Y T G D Q M I L D G P H R G D L R R A R A G A N I V P N S T G A A K A I G L V I P E I N K L D G A A Q R V P V P T 263
 QY 241 G S V T E L V A V L E K N V T V E V N A A K A A S N D S F G T E D P I V S S D I V G S G L F D A T Q T K V 300
 DB 264 G S V T E L V A V L E K N V T V E V N A A K A A S N D S F G T E D P I V S S D I V G S G L F D A T Q T K V 323
 QY 301 E V D S Q L V K V S W Y D N E M S Y T A Q L V R T L E Y F A K I A K 336
 DB 324 D V D Q K Q L V K V S W Y D N E M S Y T A Q L V R T L E Y F A K I A K 359

RESULT 9
 ID Q8KHG1 PRELIMINARY; PRT; 336 AA.
 AC Q8KHG1;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)

DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Glyceraldehyde-3-phosphate dehydrogenase.
 OS Streptococcus agalactiae, and
 OS Streptococcus infant.
 OC Bacteria; Firmicutes; Iactobacillales; Streptococcaceae;
 OC Streptococcus.
 NCBI_TaxID=1311, 1346;
 RN
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.agalactiae, and S.infant; STRAIN=ATCC 27541, and 9117;
 RA Fontaine M.C., Perez-Casal J., Song X.-W., Shellford J., Wallison P.J.,
 RA Pocher A.A.;
 RT "Immunization of dairy cattle with recombinant Gapc and chimeric CAMP
 RT antigens confers protection against heterologous challenge with
 RT Streptococcus uberis.";
 RT Submitted (Sep-2001) to the EMBL/GenBank/DBD databases.
 CC -1- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
 DEHYDROGENASE FAMILY.
 DR EMBL: AF421899; AAM73770.1; -.
 DR EMBL: AF421902; AAM73773.1; -.
 DR GO: GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.
 DR GO: GO:0016491; F:oxidoreductase activity; IEA.
 DR GO: GO:0006096; P:glycolysis; IEA.
 DR InterPro: IPR006424; GAPDH-1.
 DR InterPro: IPR00173; GAP_dhrogenase.
 DR Pfam: PF02800; gpdh. C. 1.
 DR PRINTS: PR00076; G3PDHGNASE.
 DR TIGRFAMs: TIGR01534; GAPDH-1; 1.
 DR PROSITE: PS00071; GAPDH. 1.
 DR Oxidoreductase.
 SQ SEQUENCE 336 AA; 35723 MW; ABA9E14F3ED111 CRC64;

Query Match 92.4%; Score 1584; DB 2; Length 336;
 Best Local Similarity 90.5%; Pred. No. 3.4e-98;
 Matches 304; Conservative 19; Mismatches 13; Indels 0; Gaps 0;

QY 1 M V V K G I N G F R I G R I A F R I O N V G E V T R I N D L T D P M L A H L K Y D T T G R F D G T V E V 60
 DB 1 M V V K V I N G F R I G R I A F R I O N V G E V T R I N D L T D P M L A H L K Y D T T G R F D G T V E V 60
 QY 61 K E G F E V N G N F I K V S A R D E N I M A T D G V E I V L E A T G F P A K K A A E K H L H A N G A K K V I 120
 DB 61 K O G F E V N G S F V K S A R E P A N I D M A T D G V I L E A T G F P A K K A A A O H H A N G A K K V I 120
 QY 121 T A P G S D V K V V N T N H D I L D G T E V I S G A S C T T N C L A P M A K A L D A P G I O K G I M T T H A 180
 DB 121 T A P G S D V K V V N T N H D I L D G T E V I S G A S C T T N C L A P M A K A L D A P G I O K G I M T T H A 180
 QY 181 Y T G D Q M I L D G P H R G D L R R A R A G A N I V P N S T G A A K A I G L V I P E I N K L D G A A Q R V P V P T 240
 DB 181 Y T G D Q M V I D G P H R G D L R R A R A G A N I V P N S T G A A K A I G L V I P E I N K L D G A A Q R V P V P T 240
 QY 241 G S V T E L V A V L E K N V T V E V N A A K A A S N D S F G T E D P I V S S D I V G S G L F D A T Q T K V 300
 DB 241 G S V T E L V A V L E K N T V E E I N A A K A A S N D S F G T E D P I V S S D I V G S G L F D A T Q T K V 300
 QY 301 E V D S Q L V K V S W Y D N E M S Y T A Q L V R T L E Y F A K I A K 336
 DB 301 T V D S Q L V K V S W Y D N E M S Y T A Q L V R T L E Y F A K I A K 336

RESULT 10
 ID Q8DVV3 PRELIMINARY; PRT; 337 AA.
 AC Q8DVV3;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Extracellular glyceraldehyde-3-phosphate dehydrogenase [EC
 1.2.1.12]
 GN GAPC OR SMU.360.

CS Streptococcus mutans. Lactobacillales; Streptococcaceae;
CC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxId=1309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UA159 / ATCC 700610 / Serotype C;
RX MEDLINE=22295063; PubMed=12397186;
RA Ajdic D., McShan M.M., McLaughlin R.E., Savic G., Chang J.,
RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Jin S., Qian Y.,
RA Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.;
RA Genome sequence of Streptococcus mutans UA159, a cariogenic dental
RT pathogen.
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
DR EMBL: A501483; AAS5818.1;
DR GO: GO:0004365; P-glyceraldehyde-3-phosphate dehydrogenase (p. . .); IEA.
DR GO: GO:0016491; P-oxidoreductase activity; IEA.
DR GO: GO:0006096; P-glycolysis; IEA.
DR InterPro: IPR006424; GAPDH-I.
DR InterPro: IPR00173; GAP_dhydrogenase.
DR Pfam: PF00044; gpdh_1.
DR Pfam: PF02800; gpdh_C_1.
DR PRINTS: PR00078; G3PDHGNASE.
DR TIGRFBMS: TIGR01534; GAPDH-I; 1.
DR PROSITE: PS00071; GAPDH; 1.
KW Oxidoreductase; Complete proteome.
SQ SEQUENCE 337 AA; 36068 MW; 42BF20365963C22 CRC64;

Query Match 92.1%; Score 1579.5; DB 16; Length 337;
Best Local Similarity 90.2%; Pred. No. 6.9e-98;
Matches 304; Conservative 18; Mismatches 14; Indels 1; Gaps 1;

QY 1 MVVKGINGFGRIGRLAERRIQWEGEVTRINDLPNNMLAHLLKDTTQGRDGTVEV 60
DB 1 MVVKGINGFGRIGRLAERRIQWEGEVTRINDLPNNMLAHLLKDTTQGRDGTVEV 60
QY 61 KEGFEVNGNFIKVSAREDPENIDATGVEIVLEATGFPAKKAERKLHNAKAKV 119
DB 61 KEGFEVNGNFIKVSAREDPENIDATGVEIVLEATGFPAKKAERKLHNAKAKV 120
QY 120 TAPGSDVKTVEFNTNHDLDGTEVTSASCTTCLAPAKALHDAFCIGKMTTTH 179
DB 121 TAPGSDVKTVEFNTNHDLDGTEVTSASCTTCLAPAKALHDAFCIGKMTTTH 180
QY 180 ATYGOMTIDGPRGGDLRRARAGANNIVNSTGAAKAIGVTEPLNGKLDGAARVPVP 239
DB 181 ATYGOMTIDGPRGGDLRRARAGANNIVNSTGAAKAIGVTEPLNGKLDGAARVPVP 240
QY 240 TGSVTELVAVLDKRYTVDENNAAMAAANESYGTEDPIVSSDIVGMSFSLDATTQTKV 300
DB 241 TGSVTELVAVLDKRYTVDENNAAMAAANESYGTEDPIVSSDIVGMSFSLDATTQTKV 300
QY 300 MEVDGSQLVKNVSWYDNEMSYTAQVLTLETFPAKIAK 336
DB 301 MEVDGSQLVKNVSWYDNEMSYTAQVLTLETFPAKIAK 337

RESULT 11

ID Q8KVU5 PRELIMINARY; PRT; 336 AA.
AC Q8KVU5;
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Glyceraldehyde-3-phosphate dehydrogenase.
OS Streptococcus parvulus.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxId=1348;
RN [1]
RP SEQUENCE FROM N.A.
RA Fontaine M.C., Perez-Casal J., Song X.-M., Sheldford J., Willson P.J.,
RA Potter A.A.;

RT Immunization of dairy cattle with recombinant GapC and chimeric CAMP
RT antigens confers protection against heterologous challenge with
RT Streptococcus uberis.
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
CC DEHYDROGENASE FAMILY.
DR EMBL: AF421901; AAM73772.1;
DR GO: GO:0004365; P-glyceraldehyde-3-phosphate dehydrogenase (p. . .); IEA.
DR GO: GO:0016491; P-oxidoreductase activity; IEA.
DR GO: GO:0006096; P-glycolysis; IEA.
DR InterPro: IPR006424; GAPDH-I.
DR InterPro: IPR00173; GAP_dhydrogenase.
DR Pfam: PF00044; gpdh_1.
DR Pfam: PF02800; gpdh_C_1.
DR PRINTS: PR00078; G3PDHGNASE.
DR TIGRFBMS: TIGR01534; GAPDH-I; 1.
DR PROSITE: PS00071; GAPDH; 1.
KW Oxidoreductase.
SQ SEQUENCE 336 AA; 36037 MW; 311647C25489AC9E CRC64;

Query Match 92.0%; Score 1578; DB 2; Length 336;
Best Local Similarity 90.8%; Pred. No. 8.6e-98;
Matches 305; Conservative 16; Mismatches 15; Indels 0; Gaps 0;

QY 1 MVVKGINGFGRIGRLAERRIQWEGEVTRINDLPNNMLAHLLKDTTQGRDGTVEV 60
DB 1 MVVKGINGFGRIGRLAERRIQWEGEVTRINDLPNNMLAHLLKDTTQGRDGTVEV 60
QY 61 KEGFEVNGNFIKVSAREDPENIDATGVEIVLEATGFPAKKAERKLHNAKAKV 120
DB 61 KEGFEVNGNFIKVSAREDPENIDATGVEIVLEATGFPAKKAERKLHNAKAKV 120
QY 121 TAPGSDVKTVEFNTNHDLDGTEVTSASCTTCLAPAKALHDAFCIGKMTTTH 180
DB 121 TAPGSDVKTVEFNTNHDLDGTEVTSASCTTCLAPAKALHDAFCIGKMTTTH 180
QY 181 YTGDMTIDGPRGGDLRRARAGANNIVNSTGAAKAIGVTEPLNGKLDGAARVPVP 240
DB 181 YTGDMTIDGPRGGDLRRARAGANNIVNSTGAAKAIGVTEPLNGKLDGAARVPVP 240
QY 241 GSVTELVAVLTLETFPAKIAK 336
DB 241 GSVTELVAVLTLETFPAKIAK 336
QY 301 EVDGSQLVKNVSWYDNEMSYTAQVLTLETFPAKIAK 336
DB 301 EVDGSQLVKNVSWYDNEMSYTAQVLTLETFPAKIAK 336

RESULT 12

ID Q8VVB9 PRELIMINARY; PRT; 336 AA.
AC Q8VVB9;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Glyceraldehyde-3-phosphate dehydrogenase.
GN GAPDH.
OS Streptococcus thermophilus.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxId=1108;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LMG1831;
RA van den Bogard P.T.C., Kleerebezem M., Hols P., Crispie F.,
RA Kuipers O.P., de Vos W.M.;
RA "Modulation of glycolysis by lactose availability in Streptococcus
RT thermophilus."
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
CC DEHYDROGENASE FAMILY.
DR EMBL: AF442551; AAL35177.1;

DR HSP; P03354; 38PD.
 DR GO; GO:0004365; F:glyceralddehyde-3-phosphate dehydrogenase (p. . .); IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0006096; P:glycolysis; IEA.
 DR InterPro; IPR006424; GAPDH-1.
 DR InterPro; IPR000173; GAP_dhrogenase.
 DR Pfam; PF00044; gpdh.1.
 DR Pfam; PF02800; gpdh.C.1.
 DR PRINTS; PR00078; G3PDHGNASE.
 DR TIGRfams; TIGR01534; GAPDH-1; 1.
 DR PROSITE; PS00071; GAPDH; 1.
 DR Oxidoreductase.
 SQ SEQUENCE 336 AA; 36026 MW; 52C1F25F3A7E0230 CRC64;

Query Match 91.7%; Score 1572; DB 2; Length 336;
 Best Local Similarity 90.2%; Pred. No. 2.2e-97;
 Matches 303; Conservative 18; Mismatches 15; Indels 0; Gaps 0;

QY 1 WYKVGNGFGRIGLAFRIQVGEVTRINDLDPNMLAHLKKTDTGPRDGTVEV 60
 DB 1 WYKVGNGFGRIGLAFRIQVGEVTRINDLDPNMLAHLKKTDTGPRDGTVEV 60
 QY 61 KEGFEVNGFIVSARDEPNIDWATDGEIVLBAFGPKKAAEKHLANAKKVI 120
 DB 61 KEGFEVNGFIVSARDEPNIDWATDGEIVLBAFGPKKAAEKHLANAKKVI 120
 QY 121 TAPGANTVTVFNTNHDLDGTETVSSASCTTNCCLAPMAKALHAPGJOKGIMTTIHA 180
 DB 121 TAPGANTVTVFNTNHDLDGTETVSSASCTTNCCLAPMAKALHAPGJOKGIMTTIHA 180
 QY 181 YTGDMILDGPRRGDLRRARAGANIVNSTGAAKAIGVPELNGKLDGAQRVPVPT 240
 DB 181 YTGDMILDGPRRGDLRRARAGANIVNSTGAAKAIGVPELNGKLDGAQRVPVPT 240
 QY 241 GSYVELVVTLDKVSVDLEINAAKASNDSCFTGTEPRIVSSDIWVSYSGLDPAQTQVM 300
 DB 241 GSYVELVVTLDKVSVDLEINAAKASNDSCFTGTEPRIVSSDIWVSYSGLDPAQTQVM 300
 QY 301 EYDGSOLVYKVSWYDNEYSYTAOLVTLTEYFAKIAK 336
 DB 301 EYDGSOLVYKVSWYDNEYSYTAOLVTLTEYFAKIAK 336
 DB 301 DYDGKOLVYKVSWYDNEYSYTSOLVTLTEYFAKIAK 336

RESULT 13
 Q9LSX6 PRELIMINARY; PRT; 320 AA.
 AC Q9LSX6;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Extracellular glyceralddehyde-3-phosphate dehydrogenase (Fragment).
 OS Streptococcus gordonii.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1302;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Nelson D., Goldstein J.W., Boatright K., Hartly D.W.S., Cook S.L.,
 RA Hickman P.J., Potempa J., Travis J., Mayo J.A.;
 RT "Purification and characterization of an extracellular
 glyceralddehyde-3-phosphate dehydrogenase from Streptococcus sanguis
 and cloning of the gene encoding this enzyme."
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
 DEHYDROGENASE FAMILY.
 CC EMBL; AF247678; AAF64063.1; -.
 DR HSP; P03362; 1GD1.
 DR GO; GO:0004365; F:glyceralddehyde-3-phosphate dehydrogenase (p. . .); IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0006096; P:glycolysis; IEA.
 DR InterPro; IPR006424; GAPDH-1.
 DR InterPro; IPR000173; GAP_dhrogenase.

DR Pfam; PF00044; gpdh.1.
 DR Pfam; PF02800; gpdh.C.1.
 DR PRINTS; PR00078; G3PDHGNASE.
 DR TIGRfams; TIGR01534; GAPDH-1; 1.
 DR PROSITE; PS00071; GAPDH; 1.
 DR Oxidoreductase.
 FT MON TER 1 1
 FT NON TER 320 320
 SQ SEQUENCE 320 AA; 33968 MW; 6FEC18BB01E91F0 CRC64;
 Query Match 88.5%; Score 1518; DB 2; Length 320;
 Best Local Similarity 91.2%; Pred. No. 8.4e-94;
 Matches 292; Conservative 14; Mismatches 14; Indels 0; Gaps 0;

QY 4 KYGNGFGRIGLAFRIQVGEVTRINDLDPNMLAHLKKTDTGPRDGTVEV 63
 DB 1 KYGNGFGRIGLAFRIQVGEVTRINDLDPNMLAHLKKTDTGPRDGTVEV 60
 QY 64 GFEVNGFIVSARDEPNIDWATDGEIVLBAFGPKKAAEKHLANAKKVI 123
 DB 61 GFEVNGFIVSARDEPNIDWATDGEIVLBAFGPKKAAEKHLANAKKVI 120
 QY 124 GNDVTVTVFNTNHDLDGTETVSSASCTTNCCLAPMAKALHAPGJOKGIMTTIHA 183
 DB 121 GSDVTVTVFNTNHDLDGTETVSSASCTTNCCLAPMAKALHAPGJOKGIMTTIHA 180
 QY 184 DQMLIDGPRRGDLRRARAGANIVNSTGAAKAIGVPELNGKLDGAQRVPVPT 243
 DB 181 DQMLIDGPRRGDLRRARAGANIVNSTGAAKAIGVPELNGKLDGAQRVPVPT 240
 QY 244 TELVTLDNVSVDEINAAKASNDSCFTGTEPRIVSSDIWVSYSGLDPAQTQVM 303
 DB 241 TELVTLDNVSVDEINAAKASNDSCFTGTEPRIVSSDIWVSYSGLDPAQTQVM 300
 QY 304 GSQLVKVSWYDNEYSYTAQ 323
 DB 301 GSQLVKVSWYDNEYSYTAQ 320

RESULT 14
 Q9AJT7 PRELIMINARY; PRT; 309 AA.
 AC Q9AJT7;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Glyceralddehyde-3-phosphate dehydrogenase (BC 1.2.1.12) (Fragment).
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=Clinical isolate;
 RA Amesaga M.R., Carter P.E., McKenzie H.;
 RT "Sequence variation in mef(A) and gap genes in M-phenotype
 erythromycin resistant Streptococcus pneumoniae."
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
 DEHYDROGENASE FAMILY.
 CC EMBL; AJ292048; CAC27448.1; -.
 DR HSP; P03362; 1GD1.
 DR GO; GO:0004365; F:glyceralddehyde-3-phosphate dehydrogenase (p. . .); IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0006096; P:glycolysis; IEA.
 DR InterPro; IPR006424; GAPDH-1.
 DR InterPro; IPR000173; GAP_dhrogenase.
 DR Pfam; PF00044; gpdh.1.
 DR Pfam; PF02800; gpdh.C.1.
 DR PRINTS; PR00078; G3PDHGNASE.
 DR TIGRfams; TIGR01534; GAPDH-1; 1.

DR PROSITE: PS00071; GAPDH; 1.
 KM Oxidoreductase.
 FT NON_TER 1
 FT NON_TER 309
 SQ SEQUENCE 309 AA; 32930 MW; 89E45C3BEA6A0528 CRC64;

Query Match 85.6%; Score 1468.5; DB 2; Length 309;
 Best Local Similarity 91.9%; Pred. No. 1.7e-90;
 Matches 285; Conservative 12; Mismatches 12; Indels 1; Gaps 1;

QY 9 GFGRIGLAERLQVGEVETRIINDLTPNMLAHLLKVDYTGFRDGVVEKSGFEVN 68
 DB 1 GFGRIGLAERLQVGEVETRIINDLTPNMLAHLLKVDYTGFRDGVVEKSGFEVN 60
 QY 69 GNFIVASERDPENIDMATDGVETVLEATGFPFAKKAEEKHLHANGAKKVVITAPGANDV 128
 DB 61 GNFIVASERDPENIDMATDGVETVLEATGFPFAKKAEEKHLHANGAKKVVITAPGANDV 113
 QY 129 KTVFNTNHDLDGTETVTSASCTTNCLAPAKALHDAFGIQKGLMTTHAYTGDMIL 188
 DB 120 KTVFNTNHDVLTGTETVTSASCTTNCLAPAKALQDNFVGBGLMTTHAYTGDMIL 179
 QY 189 DGFHGGDLRRARAGANIVPNSGAKALGLVPELNGKLDGSAQRVPTGSTEVA 248
 DB 180 DGFHGGDLRRARAGANIVPNSGAKALGLVPELNGKLDGSAQRVPTGSTEVA 239
 QY 249 TLDKNVSDEINAAKASNSFGYTEDPVSSDIYVSYGSLFDATQTKMEYDSGLV 308
 DB 240 VLEKNTVDEVNAAKKAASNSYGTEDPIVSSDIYVSYGSLFDATQTKVLDVGRQLV 299
 QY 309 KVVSWYDNEM 318
 DB 300 KVVSWYDNEM 309

RESULT 15

Q9AJT9 PRELIMINARY; PRT; 309 AA.

AC Q9AJT9; 01-JUN-2001 (Tremblrel. 17, Created)
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE Glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12) (Fragment).
 IE GAP.
 GX Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=clinical isolate;
 RA Amezcaga M.R., Carter P.E., Cash P., McKenzie H.;
 RT "Sequence variation in mef(A) and gap genes in M-phenotype erythromycin resistant Streptococcus pneumoniae."
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE FAMILY.
 CC EMBL: AJ292046; CAC7446.1; -.
 DR HSESP; P00362; IGD1.
 DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. .; IBA.
 DR GO; GO:0016491; F:oxidoreductase activity; IBA.
 DR GO; GO:0006096; P:glycolysis; IBA.
 DR InterPro; IPR006424; GAP_dhrcgenase.
 DR InterPro; IPR001173; GAP_dhrcgenase.
 DR Pfam; PF00044; spdh; 1.
 DR Pfam; PF02800; spdh; 1.
 DR PRINTS; PR00078; GAPDHGNASE.
 DR TIGRZAMS; TIGR01534; GAPDH-1; 1.
 DR PROSITE: PS00071; GAPDH; 1.
 KM Oxidoreductase.
 FT NON_TER 1
 FT NON_TER 309

SQ SEQUENCE 309 AA; 32916 MW; 39E9CC3BEA6F8959 CRC64;

Query Match 85.6%; Score 1467.5; DB 2; Length 309;
 Best Local Similarity 91.6%; Pred. No. 1.3e-90;
 Matches 284; Conservative 13; Mismatches 12; Indels 1; Gaps 1;

QY 9 GFGRIGLAERLQVGEVETRIINDLTPNMLAHLLKVDYTGFRDGVVEKSGFEVN 68
 DB 1 GFGRIGLAERLQVGEVETRIINDLTPNMLAHLLKVDYTGFRDGVVEKSGFEVN 60
 QY 69 GNFIVASERDPENIDMATDGVETVLEATGFPFAKKAEEKHLHANGAKKVVITAPGANDV 128
 DB 61 GNFIVASERDPENIDMATDGVETVLEATGFPFAKKAEEKHLHANGAKKVVITAPGANDV 119
 QY 129 KTVFNTNHDLDGTETVTSASCTTNCLAPAKALHDAFGIQKGLMTTHAYTGDMIL 188
 DB 120 KTVFNTNHDVLTGTETVTSASCTTNCLAPAKALQDNFVGBGLMTTHAYTGDMIL 179
 QY 189 DGFHGGDLRRARAGANIVPNSGAKALGLVPELNGKLDGSAQRVPTGSTEVA 248
 DB 180 DGFHGGDLRRARAGANIVPNSGAKALGLVPELNGKLDGSAQRVPTGSTEVA 239
 QY 249 TLDKNVSDEINAAKASNSFGYTEDPVSSDIYVSYGSLFDATQTKMEYDSGLV 308
 DB 240 VLEKNTVDEVNAAKKAASNSYGTEDPIVSSDIYVSYGSLFDATQTKVLDVGRQLV 299
 QY 309 KVVSWYDNEM 318
 DB 300 KVVSWYDNEM 309

Search completed: May 5, 2004, 14:21:09
 Job time : 46 secs

1	1775	100.0	336	5	AAAS06065	AAAS06065	Streptococcus
2	1775	100.0	336	5	AAAS06063	AAAS06063	Streptococcus
3	1714	99.9	336	5	ABP29960	ABP29960	Streptococcus
4	1714	99.9	336	6	ABH46655	ABH46655	Protein e
5	1714	99.9	345	5	ABP29106	ABP29106	Streptococcus
6	1711	99.8	336	2	AAH56486	AAH56486	Plasman
7	1711	99.8	336	4	AAH56581	AAH56581	Streptococcus
8	1656.5	96.6	448	5	AAAS06064	AAAS06064	Streptococcus
9	1618	94.3	336	5	AAAS06067	AAAS06067	Streptococcus
10	1618	94.3	336	5	AAAS06041	AAAS06041	Streptococcus
11	1609	93.8	336	5	ABP30758	ABP30758	Streptococcus
12	1602	93.4	336	5	AAAS06066	AAAS06066	Streptococcus
13	1602	93.4	336	5	AAAS06040	AAAS06040	Streptococcus
14	1593.5	92.9	335	6	ABH02516	ABH02516	S. pneumoniae
15	1593.5	92.9	359	4	AAH37576	AAH37576	Streptococcus
16	1593.5	92.9	359	6	ABH46652	ABH46652	Protein e
17	1592.5	92.9	339	4	AAH38000	AAH38000	Streptococcus
18	1584	92.4	336	5	AAAS06069	AAAS06069	Streptococcus
19	1584	92.4	336	5	AAAS06043	AAAS06043	Streptococcus
20	1579.5	92.1	337	6	ABH344330	ABH344330	Protein e
21	1578	92.0	336	5	AAAS06068	AAAS06068	Streptococcus
22	1578	92.0	336	5	AAAS06042	AAAS06042	Streptococcus
23	1549.5	90.3	333	2	AAAS05089	AAAS05089	Streptococcus
24	1549.5	90.3	333	5	ABH54583	ABH54583	S. pneumoniae
25	1549.5	90.3	333	7	ADCA51535	ADCA51535	S. pneumoniae

26	1410	82.2	336	5	AB355601	Lactococc
27	1337.5	78.0	333	4	AAU55284	Aa355254
28	1337.5	78.0	333	6	ABU29370	Batreroccc
29	1333.5	77.8	336	5	AB53868	Abb29370
30	1325.5	77.3	333	7	AD295182	Abb53868
31	1277	74.5	332	7	ABU25342	Ad95142
32	1262	74.5	335	6	ABU25342	F. faeciu
33	1262	73.6	336	5	AB48810	Abb48810
34	1255.5	73.6	336	6	ABU2558	Abb48810
35	1217.5	71.0	334	6	ABU23595	Abb423595
36	1217.5	69.8	335	6	ABU43810	Protein e
37	1196.5	69.8	334	6	ABU37737	Protein e
38	1196.5	69.8	357	6	ABP81005	Abp81005
39	1190	69.8	357	6	ABU37077	N. gonorr
40	1190	69.4	336	6	AA683058	Abu37077
41	1190	69.4	348	5	ABU42966	S. epider
42	1183	69.0	336	6	ABP40668	Aa42966
43	1183	69.0	336	6	AAU66977	Abp40668
44	1183	69.0	336	6	ABU4122	Staphyloc
45	1183	69.0	336	6	ABU19195	Aa4122
					AbU19195	Staphyloc
					ABU16073	Pathogen
						AbU16073
						Protein e

XX	RESULT 1
AA	AAAM50665
ID	AAAM50665 standard; protein; 336 AA.
XX	
AC	AAAM50665;
XX	
DT	08-APR-2002 (first entry)
XX	
DE	Streptococcus dysgalactiae gapC plasmin binding protein DysgapC.
XX	
XX	DysGapC; GapC; plasmin binding protein; epitope; infection; vaccine;
KW	immunisation; mastitis; therapy.
OS	Streptococcus dysgalactiae.
XX	
XX	MO200196379-A2.
XX	
XX	20-DEC-2001.
PD	
XX	
PF	11-JUN-2001; 2001WO-CA00836.
XX	
XX	12-JUN-2000; 2000US-0211247P.
PR	
XX	(UYSA-) UNIV SASKATCHEWAN.
PA	
XX	Potter AA, Perez-Casal J, Fontaine M;
P1	
DR	WPI; 2002-038051/13.
XX	
XX	N-PSDB; AAM91328.
XX	
XX	Novel GapC multiple epitope fusion polypeptide comprising antigenic
PT	determinant of Streptococcus dysgalactiae, S. agalactiae, S. uberis, S.
PT	parvberis, or S. iniae GapC protein useful for treating mastitis in
P1	vertebrates.
XX	
XX	Claim 8; Fig 1A-B; 116pp; English.
XX	
XX	The present sequence is that of the GapC plasmin binding protein,
CC	DysgapC, of Streptococcus dysgalactiae ATCC 43078, an isolate from a case
CC	of bovine mastitis. The invention relates to novel GapC multiple epitope
CC	fusion proteins that comprise epitopes from 1 or more of Streptococcus
CC	dysgalactiae, Streptococcus agalactiae, Streptococcus parvberis,
CC	Streptococcus uberis and Streptococcus iniae (see AAM50665-69). A claimed
CC	example is PolyGap4 (see AAM50664). Expression vectors and host cells for
CC	production of the multiple epitope fusion proteins are provided. The
CC	multiple epitope proteins are used in claimed vaccines, especially a
CC	preventing a bacterial infection in a vertebrate, especially also
CC	streptococcal infection, and particularly mastitis. They are also used in
CC	

CC claimed methods of detecting Streptococcus antibodies. The multiple
CC epitope proteins are capable of eliciting broad immunity against a
CC variety of streptococcal infections while minimising the number of
CC antigens present in the final formulation and concomitantly reducing
CC production costs

XX Sequence 336 AA;

Query Match 100.0%; Score 1715; DB 5; Length 336;
Best Local Similarity 100.0%; Pred. No. 8.6e-145;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVTGNGNGRIGRLAFRRIONVEGVETRLINDLTPNMLAHLKXDTTQGRFDGTVFV 60
DB 1 MVTGNGNGRIGRLAFRRIONVEGVETRLINDLTPNMLAHLKXDTTQGRFDGTVFV 60
QY 61 KEGGFVNGNFIKVSARBDPENIDMTGVEIVLAEATGFPKKEAAEKHLHANGAKKVI 120
DB 61 KEGGFVNGNFIKVSARBDPENIDMTGVEIVLAEATGFPKKEAAEKHLHANGAKKVI 120
QY 121 TAPGNDVKTVEFNTNHDILDGTEVTSASCTTCLAPAKALHDAFGIOGMLTTIHA 180
DB 121 TAPGNDVKTVEFNTNHDILDGTEVTSASCTTCLAPAKALHDAFGIOGMLTTIHA 180
QY 181 YTGDDMILDPHRRGDLRRARAGANIVPNSGAAKAIGLVTPELNGKLDGAARVPVPT 240
DB 181 YTGDDMILDPHRRGDLRRARAGANIVPNSGAAKAIGLVTPELNGKLDGAARVPVPT 240
QY 241 GSYTELVTTLDKRVSDVINAMKAASNDSTGTEDPYSSDIIVGSYGLFDAQTKVM 300
DB 241 GSYTELVTTLDKRVSDVINAMKAASNDSTGTEDPYSSDIIVGSYGLFDAQTKVM 300
QY 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTEYFAKIAK 336
DB 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTEYFAKIAK 336

RESULT 2

ID AAM50639 standard; protein; 336 AA.

XX AAM50639;

DI 04-APR-2002 (first entry)

DE Streptococcus dysgalactiae gapc plasmin binding protein.

KW GapC; plasmin-binding protein; DysgalGapC; infection; mastitis; vaccine;
diagnosis; therapy.

CS Streptococcus dysgalactiae.

PN MO200196381-A2.

PD 20-DEC-2001.

PF 11-JUN-2001; 2001MO-CA000838.

PR 12-JUN-2000; 2000US-0211022P.

XX (UTSA-) UNIV SASKATCHEWAN.

PI Bolton AJ, Perez-Casal J, Fontaine M, Potter AA;

DR MPI; 2002-130725/17.

DR N-EBDB; ABA91248.

PT Novel isolated GapC protein of Streptococcus dysgalactiae, S. agalactiae,
S. uberis, S. parvulus, or S. infantiae, useful as vaccine component for
treating streptococcal infection which causes mastitis in vertebrates.

PS Claim 1(a); Fig 1A-B; 107pp; English.

CC The present sequence is that of the GapC plasmin binding protein
CC (DysgalGapC) of Streptococcus dysgalactiae ATCC 43078, a clinical isolate
CC from bovine mastitis. It is encoded by the GapC gene given in ABA91248.
CC GapC protein, which has no signal sequence or membrane anchor domain, is
CC capable of eliciting an immune response in a vertebrate. The invention
CC provides the GapC genes and proteins of 5 Streptococcus species, as well
CC as recombinant vectors, host cells and vaccine compositions comprising
CC GapC polynucleotides or proteins. The vaccines are used to treat or
CC prevent a bacterial infection, especially a streptococcal infection, and
CC mastitis in particular (claimed). GapC proteins are also used in claimed
CC methods for detecting GapC antibodies, and to raise antibodies that are
CC used in claimed methods for detecting GapC proteins

XX Sequence 336 AA;

Query Match 100.0%; Score 1715; DB 5; Length 336;
Best Local Similarity 100.0%; Pred. No. 8.6e-145;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVTGNGNGRIGRLAFRRIONVEGVETRLINDLTPNMLAHLKXDTTQGRFDGTVFV 60
DB 1 MVTGNGNGRIGRLAFRRIONVEGVETRLINDLTPNMLAHLKXDTTQGRFDGTVFV 60
QY 61 KEGGFVNGNFIKVSARBDPENIDMTGVEIVLAEATGFPKKEAAEKHLHANGAKKVI 120
DB 61 KEGGFVNGNFIKVSARBDPENIDMTGVEIVLAEATGFPKKEAAEKHLHANGAKKVI 120
QY 121 TAPGNDVKTVEFNTNHDILDGTEVTSASCTTCLAPAKALHDAFGIOGMLTTIHA 180
DB 121 TAPGNDVKTVEFNTNHDILDGTEVTSASCTTCLAPAKALHDAFGIOGMLTTIHA 180
QY 181 YTGDDMILDPHRRGDLRRARAGANIVPNSGAAKAIGLVTPELNGKLDGAARVPVPT 240
DB 181 YTGDDMILDPHRRGDLRRARAGANIVPNSGAAKAIGLVTPELNGKLDGAARVPVPT 240
QY 241 GSYTELVTTLDKRVSDVINAMKAASNDSTGTEDPYSSDIIVGSYGLFDAQTKVM 300
DB 241 GSYTELVTTLDKRVSDVINAMKAASNDSTGTEDPYSSDIIVGSYGLFDAQTKVM 300
QY 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTEYFAKIAK 336
DB 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTEYFAKIAK 336

RESULT 3

ID ABP29960 standard; protein; 336 AA.

XX ABP29960;

DI 02-JUL-2002 (first entry)

DE Streptococcus polypeptide SEQ ID NO 9036.

KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
group A streptococcus; Streptococcus pyogenes; antibacterial;
antiinflammatory; infection; vaccine; meningitis; gene therapy.

CS Streptococcus pyogenes.

PN MO200234771-A2.

PD 02-MAY-2002.

PF 23-OCT-2001; 2001MO-GB004789.

PR 27-OCT-2000; 2000GB-00026333.

PR 24-NOV-2000; 2000GB-00028727.

PR 07-MAR-2001; 2001GB-00005640.

PA (CHIR-) CHIRON SPA.
(GENO-) INST GENOMIC RES.

PI Telford J, Masignani V, Margarit Y Rosl, Grandi G, Frazer C;
PI Tetelin H;
XX
XX
DR WPI; 2002-352536/38.
DR N-PSDB; AEN70591.
XX
XX
PT New Streptococcus protein for the treatment or prevention of infection or
PT disease caused by Streptococcus bacteria, such as meningitis, and for
PI detecting a compound that binds to the protein.
XX
XX
PS Claim 1; Page 4031; 4525pp; English.
XX
CC The invention relates to a protein (ABP25413-ABP20893) from group B
CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
CC the specification. The proteins have antibacterial and anti-inflammatory
CC activity. (1), nucleic acid encoding (1), ABN6004-ABN71526 and
CC antibodies that bind (1) are used in the manufacture of medicaments for
CC the treatment or prevention of infection or disease caused by
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
CC Nucleic acid encoding (1) is used to detect Streptococcus in a
CC biological sample. (1) is used to determine whether a compound binds to
CC (1). A composition comprising (1) or a nucleic acid encoding (1), may be
CC used as a vaccine or diagnostic composition. The disease caused by
CC Streptococcus that is prevented or treated may be meningitis. Nucleic
CC acid encoding (1) may be used to recombinantly produce (1) and may be
CC used in gene therapy. Antibodies to (1) are used for affinity
CC chromatography, immunoassays, and distinguishing/identifying
CC Streptococcus proteins
XX
SQ Sequence 336 AA;
Query Match 99.9%; Score 1714; DB 5; Length 336;
Best Local Similarity 99.7%; Pred. No. 1,1e-144;
Matches 335; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 MYVKGINGFGRIGRLAERIONVEGVEVRINDLTPDMLAHLLKDTTQGRFGTVEV 60
DB 1 MYVKGINGFGRIGRLAERIONVEGVEVRINDLTPDMLAHLLKDTTQGRFGTVEV 60
QY 61 KEGGFVNGNFIKVSARDEPNIDMTDGEIVLEATGFPAKKEAKHILHANGAKKVI 120
DB 61 KEGGFVNGNFIKVSARDEPNIDMTDGEIVLEATGFPAKKEAKHILHANGAKKVI 120
QY 121 TARGNDKKTVEVTHNDLIDGTEVVISGASCTTNCLAPMAKALHDAFGIOKGLMTTHA 180
DB 121 TARGNDKKTVEVTHNDLIDGTEVVISGASCTTNCLAPMAKALHDAFGIOKGLMTTHA 180
QY 121 TARGNDKKTVEVTHNDLIDGTEVVISGASCTTNCLAPMAKALHDAFGIOKGLMTTHA 180
DB 121 TARGNDKKTVEVTHNDLIDGTEVVISGASCTTNCLAPMAKALHDAFGIOKGLMTTHA 180
QY 181 YTGDQWILDGHRGGLRRARAGANIVPNTGANAAGIATPELNGKLDGAARVPVPT 240
DB 181 YTGDQWILDGHRGGLRRARAGANIVPNTGANAAGIATPELNGKLDGAARVPVPT 240
QY 241 GSVELVYVTLDDKNSVDEINAAKASNDSPGYTEDPIVSDIVSGYSLFATQTKM 300
DB 241 GSVELVYVTLDDKNSVDEINAAKASNDSPGYTEDPIVSDIVSGYSLFATQTKM 300
QY 301 EYDGSQLYKVVSWTNDNMTAQLVPTTEFPATKX 336
DB 301 EYDGSQLYKVVSWTNDNMTAQLVPTTEFPATKX 336
DB 301 EYDGSQLYKVVSWTNDNMTAQLVPTTEFPATKX 336
RESULT 4
ABU46455
ID ABU46455 standard; protein; 336 AA.
XX
XX
AC ABU46455;
XX
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by Prokaryotic essential gene #31982.
XX
XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX

OS Streptococcus pyogenes.
XX
XX
PM MO200277183-A2.
XX
XX
PD 03-OCT-2002.
XX
XX
PE 21-MAR-2002; 2002WC-US090107.
XX
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
XX
PA (ELIT-) ELITRA PHARM INC.
XX
XX
PI Wang J, Zamudio C, Malone C, Haeelbeck R, Ohlsen KU, Zvejdik JW,
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
XX
DR WPI: 2003-029926/02.
DR N-PSDB; ACA50325.
XX
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
XX
PS Claim 25; SEQ ID NO 74379; 1766pp; English.
XX
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 623 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-regulated gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pat_sequences
XX
XX
SQ Sequence 336 AA;
Query Match 99.9%; Score 1714; DB 6; Length 336;
Best Local Similarity 99.7%; Pred. No. 1,1e-144;
Matches 335; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 MYVKGINGFGRIGRLAERIONVEGVEVRINDLTPDMLAHLLKDTTQGRFGTVEV 60
DB 1 MYVKGINGFGRIGRLAERIONVEGVEVRINDLTPDMLAHLLKDTTQGRFGTVEV 60
QY 61 KEGGFVNGNFIKVSARDEPNIDMTDGEIVLEATGFPAKKEAKHILHANGAKKVI 120
DB 61 KEGGFVNGNFIKVSARDEPNIDMTDGEIVLEATGFPAKKEAKHILHANGAKKVI 120

QY 121 TARGNDVKTWVNTNHDILDGETEVIISGASCTTNCLAPMAKALHDAFGIOGLMTTTHA 180
 DB 121 TARGNDVKTWVNTNHDILDGETEVIISGASCTTNCLAPMAKALHDAFGIOGLMTTTHA 180
 QY 181 YTGDMITLDGPHRGDGLRRARAGANIVPNSGAAKAGLVIPINLKGKLDGAAGRPVPT 240
 DB 181 YTGDMITLDGPHRGDGLRRARAGANIVPNSGAAKAGLVIPINLKGKLDGAAGRPVPT 240
 QY 241 GSVELVITLIDKXVSVDENINAAKKAASNDSPGYTEDPIVSSDIVGSYGLSDPATQTKM 300
 DB 241 GSVELVITLIDKXVSVDENINAAKKAASNDSPGYTEDPIVSSDIVGSYGLSDPATQTKM 300
 QY 301 EVDGSQLVKXVSWYDNEMSYTAQVLTLEFPKIAK 336
 DB 301 EVDGSQLVKXVSWYDNEMSYTAQVLTLEFPKIAK 336

RESULT 5

ABP29106
 ID ABP29106 standard; protein; 345 AA.

AC ABP29106;

DT 02-JUL-2002 (first entry)

DE Streptococcus polypeptide SEQ ID NO 7388.

KM Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
 KM group A streptococcus; Streptococcus pyogenes; antibacterial;
 KM antiinflammatory; infection; vaccine; meningitis; gene therapy.
 XX Streptococcus pyogenes.

EM W0230234771-A2.

ED 02-MAY-2002.

PF 29-OCT-2001; 2001WO-GB004789.

PR 27-OCT-2000; 2000GB-00026333.

PR 24-NOV-2000; 2000GB-00028727.

PR 07-MAR-2001; 2001GB-00035640.

PA (CHIR-) CHIRON SPA.

PA (GENO-) INST GENOMIC RES.

PI Telford C, Masigiani V, Margarit Y Rosi, Grandi G, Fraser C;

PI Tettelin H;

DR MPI; 2002-352536/38.

DR N-PSDB; ABN69737.

PS Claim 1; Page 3888; 4525pp; English.

CC The invention relates to a protein (ABP25413-ABP30895) from group B
 CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
 CC the specification. The proteins have antibacterial and antiinflammatory
 CC activity. (1), nucleic acids encoding (1), ABN6044-ABN71526 and
 CC antibodies that bind (1) are used in the manufacture of medicaments for
 CC the treatment or prevention of infection or disease caused by
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
 CC Nucleic acids encoding (1) are used to detect Streptococcus in a
 CC biological sample. (1) is used to determine whether a compound binds to
 CC (1). A composition comprising (1) or a nucleic acid encoding (1), may be
 CC used as a vaccine or diagnostic composition. The disease caused by
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic
 CC acid encoding (1) may be used to recombinantly produce (1) and may be

CC used in gene therapy. Antibodies to (1) are used for affinity
 CC chromatography, immunoassays, and distinguishing/identifying
 CC Streptococcus proteins
 XX

SQ Sequence 345 AA;

Query Match 99.94; Score 1714; DB 5; Length 345;
 Best Local Similarity 99.74; Pred. No. 1,1e-144;
 Matches 335; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYKVGINFGHIGRLAFRIQINVEGVETRIINDLTDPNMLAHLKYDTTGPRDGVVEV 60
 DB 10 MYKVGINFGHIGRLAFRIQINVEGVETRIINDLTDPNMLAHLKYDTTGPRDGVVEV 69
 QY 61 KEGGFVNGNFVKYASRDPENIDMATDGEVIEVATGFPKKEAAEKHLNANAKKVVV 120
 DB 70 KEGGFVNGNFVKYASRDPENIDMATDGEVIEVATGFPKKEAAEKHLNANAKKVVV 129
 QY 121 TARGNDVKTWVNTNHDILDGETEVIISGASCTTNCLAPMAKALHDAFGIOGLMTTTHA 180
 DB 130 TARGNDVKTWVNTNHDILDGETEVIISGASCTTNCLAPMAKALHDAFGIOGLMTTTHA 189
 QY 181 YTGDMITLDGPHRGDGLRRARAGANIVPNSGAAKAGLVIPINLKGKLDGAAGRPVPT 240
 DB 190 YTGDMITLDGPHRGDGLRRARAGANIVPNSGAAKAGLVIPINLKGKLDGAAGRPVPT 249
 QY 241 GSVELVITLIDKXVSVDENINAAKKAASNDSPGYTEDPIVSSDIVGSYGLSDPATQTKM 300
 DB 250 GSVELVITLIDKXVSVDENINAAKKAASNDSPGYTEDPIVSSDIVGSYGLSDPATQTKM 309
 QY 301 EVDGSQLVKXVSWYDNEMSYTAQVLTLEFPKIAK 336
 DB 310 EVDGSQLVKXVSWYDNEMSYTAQVLTLEFPKIAK 345

RESULT 6

AAR56486
 ID AAR56486 standard; protein; 336 AA.

AC AAR56486;

DT 25-MAR-2003 (revised)

DT 22-FEB-1995 (first entry)

DE Plasmin receptor.

KM Plasmin receptor; isolate 64/14; plasmin; alpha-2-antiplasmin;
 KM plasminogen activator; bleeding; reocclusion; thrombosis;
 KM pulmonary embolism; clots.

XX Streptococcus pyogenes.

OS US5328996-A.

PN 12-JUL-1994.

PD 10-AUG-1992; 92US-00928462.

PF 29-MAR-1989; 89US-00330849.

PR 16-MAY-1990; 90US-00524411.

PA (UYEL) UNIV FLORIDA RES FOUND INC.

PI Von Mering G, Broder C, Boyle MDP, Lottenberg R;

PI MPI; 1994-225327/27.

DR N-PSDB; AAQ70705.

PS New DNA encoding bacterial plasmin receptor - useful as thrombolytic
 PT agents, used with plasminogen activator or bound to plasmin, also useful
 PT in vaccines against bacterial infection.

PS Claim 1; Col 27-30; 19pp; English.

XX This sequence represents the *S. pyogenes* plasmin receptor. The DNA
 CC encoding this sequence was isolated from the *S. pyogenes* clinical isolate
 CC 64/14. The plasmin receptor has a very high affinity for plasmin which,
 CC when bound, retains its enzymatic activity but is not regulated
 CC (inhibited) by alpha-2-antiplasmin. The receptor protein, when
 CC administered concurrently or sequentially, prolongs the activity of
 CC plasminogen activator (PA) so allows a reduction in dose, and thus lowers
 CC the risk of bleeding, and may prevent recocclusion of blood vessels. The
 CC protein may be coupled to a fibrin-specific monoclonal antibody to
 CC provide targeting to clots. The plasmin receptor may be useful in human
 CC or veterinary medicine, for treatment of thrombosis and pulmonary
 CC embolism, and for solubilizing clots in catheters or shunts. (Updated on
 CC 25-MAR-2003 to correct PF field.)
 XX

SQ Sequence 336 AA;

Query Match 99.8%; Score 1711; DB 2; Length 336;

Best Local Similarity 99.4%; Pred. No. 2e-144;

Matches 334; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVTGKINGFGRIGRLAFRIQNVGCVETRIINDLDPNMLAHLLKIDTTOGRFGDTVEV 60
 DB 1 MVTGKINGFGRIGRLAFRIQNVGCVETRIINDLDPNMLAHLLKIDTTOGRFGDTVEV 60
 QY 61 KEGFEVNGNFIVKSAERDPENIDMATDVEIVLATGFFAKKEAEKHLHANGAKKVI 120
 DB 61 KEGFEVNGNFIVKSAERDPENIDMATDVEIVLATGFFAKKEAEKHLHANGAKKVI 120
 QY 121 TAPGNDVKTVEFNTNHDILDTETIVISGASCTTNCIAPMAKALHDAFGIQKGLMTTHA 180
 DB 121 TAPGNDVKTVEFNTNHDILDTETIVISGASCTTNCIAPMAKALHDAFGIQKGLMTTHA 180
 QY 181 YTGDDMIIDGPHRGDLRRARAGANIVPNSGAAKAIGLVPELNGSLDGAARVPPT 240
 DB 181 YTGDDMIIDGPHRGDLRRARAGANIVPNSGAAKAIGLVPELNGSLDGAARVPPT 240
 QY 241 GSATLVTLLDKNVSYDEINAAKASNDSPGTEDPVSSDIIVGYSGLPDAQTQVM 300
 DB 241 GSATLVTLLDKNVSYDEINAAKASNDSPGTEDPVSSDIIVGYSGLPDAQTQVM 300
 QY 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTEYPAKIAK 336
 DB 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTEYPAKIAK 336

RESULT 7

AAV85681

ID AAV85681 standard; protein; 336 AA.

XX AAV85681;

DT 13-FEB-2001 (first entry)

XX Streptococcal plasmin receptor amino acid sequence.

XX Plasmin receptor; p1r; immune response; vaccination; antibacterial;

XX Streptococcal infection; antihelminthic.

XX Streptococcus pyogenes.

XX US6136323-A.

XX 24-OCT-2000.

XX 11-JUL-1994; 94US-00273247.

XX 29-MAR-1989; 89US-00330849.

XX 16-MAY-1990; 90US-00524411.

XX 10-AUG-1992; 92US-00928462.

XX (UTFL) UNIV FLORIDA RES FOUND INC.

P1 Von Mering G, Broder C, Lottenberg R, Boyle MDP;
 XX WPI: 2001-006210/01.
 DR N-PSNB; AAC66140.
 XX Raising an immune response in a mammal, especially for vaccination
 PT against group A streptococcal infections, comprises administering an
 PT isolated streptococcal plasmin receptor protein.
 XX

PS Claim 2; Col 27-30; 17pp; English.

XX Polynucleotide sequence AAC66140 encodes the Streptococcal plasmin
 CC receptor protein AAV85681. The protein is used in a method for raising an
 CC immune response in a mammal. The method comprises administering the
 CC plasmin receptor protein. The method is useful as a vaccination against
 CC group A Streptococcal infections and potentially against a broad range of
 CC infections associated with pathogens expressing glyceraldehyde-3-
 CC phosphate dehydrogenase (GAPDH) or GAPDH-like proteins at their cell
 CC surfaces. The vaccine has antibacterial and antihelminthic activity
 CC

SQ Sequence 336 AA;

Query Match 99.8%; Score 1711; DB 4; Length 336;

Best Local Similarity 99.4%; Pred. No. 2e-144;

Matches 334; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVTGKINGFGRIGRLAFRIQNVGCVETRIINDLDPNMLAHLLKIDTTOGRFGDTVEV 60
 DB 1 MVTGKINGFGRIGRLAFRIQNVGCVETRIINDLDPNMLAHLLKIDTTOGRFGDTVEV 60
 QY 61 KEGFEVNGNFIVKSAERDPENIDMATDVEIVLATGFFAKKEAEKHLHANGAKKVI 120
 DB 61 KEGFEVNGNFIVKSAERDPENIDMATDVEIVLATGFFAKKEAEKHLHANGAKKVI 120
 QY 121 TAPGNDVKTVEFNTNHDILDTETIVISGASCTTNCIAPMAKALHDAFGIQKGLMTTHA 180
 DB 121 TAPGNDVKTVEFNTNHDILDTETIVISGASCTTNCIAPMAKALHDAFGIQKGLMTTHA 180
 QY 181 YTGDDMIIDGPHRGDLRRARAGANIVPNSGAAKAIGLVPELNGSLDGAARVPPT 240
 DB 181 YTGDDMIIDGPHRGDLRRARAGANIVPNSGAAKAIGLVPELNGSLDGAARVPPT 240
 QY 241 GSATLVTLLDKNVSYDEINAAKASNDSPGTEDPVSSDIIVGYSGLPDAQTQVM 300
 DB 241 GSATLVTLLDKNVSYDEINAAKASNDSPGTEDPVSSDIIVGYSGLPDAQTQVM 300
 QY 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTEYPAKIAK 336
 DB 301 EVDGSQLVKVSWYDNEMSYTAQLVRLTEYPAKIAK 336

RESULT 8

AAV50664

ID AAV50664 standard; protein; 448 AA.

XX AAV50664;

DT 29-AUG-2003 (revised)

DT 08-APR-2002 (first entry)

XX Streptococcus GapC multiepitope fusion Polycap4.

XX Polycap4; GapC; plasmin binding protein; epitope; infection; vaccine;

XX immunisation; mastitis; therapy.

XX Streptococcus dysgalactiae.

XX Streptococcus agalactiae.

XX Streptococcus parauberis.

XX Chimeric.

XX WO200196379-A2.

XX 20-DEC-2001.

XX 11-JUN-2001; 2001MO-CA000836.
 PE
 XX
 ER 12-JUN-2000; 2000US-0211247P.
 XX
 XX (UYSA-) UNIV SASKATCHEWAN.
 XX
 XX PI Potter AA, Perez-Casal J, Fontaine M;
 XX DR MPI; 2002-098051/13.
 XX DR N-PSDB; ABA91327.
 XX
 FT Novel GApC multiple epitope fusion polypeptide comprising antigenic
 ER determinant of Streptococcus dysgalactiae, S. agalactiae, S. uberis, S.
 PT parabratis, or S. iniae GApC protein useful for treating mastitis in
 PT vertebrates.
 XX
 XX Claim 9; Fig 6A-C; 116pp; English.
 XX
 CC The present sequence is that of a novel multiple epitope fusion protein,
 CC designated PolyGap4, comprising the entire amino acid sequence of the
 CC Streptococcus dysgalactiae GApC plasmin binding protein in addition to
 CC unique amino acid sequences from the Streptococcus parabratis and
 CC Streptococcus agalactiae GApC proteins. The multiple epitope protein is
 CC produced in host cells transformed with an expression vector comprising a
 CC chimeric gene (see ABA91327) encoding the protein. PolyGap4 is an example
 CC of novel GApC multiple epitope fusion proteins of the invention that
 CC comprise epitopes from 1 or more of S. dysgalactiae, S. agalactiae, S.
 CC parabratis, Streptococcus uberis and Streptococcus iniae (see AAM50665-
 CC 69). The multiple epitope fusion proteins are used in claimed vaccines
 CC for treating or preventing a bacterial infection in a vertebrate,
 CC especially a streptococcal infection, and particularly mastitis. They are
 CC also used in claimed methods of detecting Streptococcus antibodies. The
 CC multiple epitope protein is capable of eliciting broad immunity against a
 CC variety of streptococcal infections while minimizing the number of
 CC antigens present in the final formulation and concomitantly reducing
 CC production costs. (Updated on 29-AUG-2003 to standardise OS field)
 CC
 XX
 XX Sequence 448 AA:
 SQ
 Query Match 96.6%; Score 1656.5; DB 5; Length 448;
 Best Local Similarity 79.6%; Pred. No. 2,26-139;
 Matches 335; Conservative 0; Mismatches 1; Indels 85; Gaps 1;
 QY 1 M V V K G I N G F G R I G R L A F R R I O N V E G V E T R I N D I D P M L A H L K Y D T T O G R F D G T V E V 60
 DB 28 M V V K G I N G F G R I G R L A F R R I O N V E G V E T R I N D I D P M L A H L K Y D T T O G R F D G T V E V 87
 QY 61 K E G G E V N G N F I K V S A R D P E ----- 81
 DB 88 K E G G E V N G N F I K V S A R D P E N I D W A T D G V E I V L A L E G T V E V K D G F D W N G K F I V S A E 147
 QY 82 ----- N I D W A T G V E I V L E 95
 DB 148 K D E Q I D W A T D G V E I V L E I D G T V E V K E G G E V N G Q F V K V S A R R P A N I D W A T D G V E I V L E 207
 QY 96 A T G F P A K E A A E K E I H A N G A K K V I T A P G A N D V T V E W T N H D I I D G T E V I S A S C T T N 155
 DB 208 A T G F P A K E A A E K E I H A N G A K K V I T A P G A N D V T V E W T N H D I I D G T E V I S A S C T T N 267
 QY 156 C I A P M A K A L H D A F G I O K G L M T T I H A T G D Q M I L D G H R G G D I R R A R A G A N I V P N T G A A 215
 DB 268 C I A P M A K A L H D A F G I O K G L M T T I H A T G D Q M I L D G H R G G D I R R A R A G A N I V P N T G A A 327
 QY 216 K A I G V I P E L I N G K L D G A A G V P P G S V T E I V T L D K N V S V E I N A A K A A S N D S G Y T E 275
 DB 328 K A I G V I P E L I N G K L D G A A G V P P G S V T E I V T L D K N V S V E I N A A K A A S N D S G Y T E 387
 QY 276 D P V S S D I V G S Y G S L P D A T O T K V E V D G S Q L V E V V S W Y D N E M S Y T A O L V T L E Y P A K I A 335
 DB 388 D P V S S D I V G S Y G S L P D A T O T K V E V D G S Q L V E V V S W Y D N E M S Y T A O L V T L E Y P A K I A 447
 QY 336 K 336

DB 448 K 448
 RESULT 9
 AAM50667
 ID AAM50667 standard; protein; 336 AA.
 XX
 XX AAM50667;
 XX
 DT 08-APR-2002 (first entry)
 XX
 DE Streptococcus uberis GApC plasmin binding protein UderGapC.
 XX
 XX UderGapC; GApC; plasmin binding protein; epitope; infection; vaccine;
 KW immunisation; mastitis; therapy.
 XX
 OS Streptococcus uberis.
 XX
 PM MO200196379-A2.
 PD 20-DEC-2001.
 XX
 PF 11-JUN-2001; 2001MO-CA000836.
 XX
 ER 12-JUN-2000; 2000US-0211247P.
 XX
 XX (UYSA-) UNIV SASKATCHEWAN.
 XX
 PI Potter AA, Perez-Casal J, Fontaine M;
 XX
 XX MPI; 2002-098051/13.
 DR N-PSDB; ABA91330.
 XX
 FT Novel GApC multiple epitope fusion polypeptide comprising antigenic
 ER determinant of Streptococcus dysgalactiae, S. agalactiae, S. uberis, S.
 PT parabratis, or S. iniae GApC protein useful for treating mastitis in
 PT vertebrates.
 XX
 XX Claim 8; Fig 3A-B; 116pp; English.
 XX
 CC The present sequence is that of the GApC plasmin binding protein,
 CC UderGapC, of Streptococcus uberis ATCC 9927. The invention relates to
 CC novel GApC multiple epitope fusion proteins that comprise epitopes from 1
 CC or more of Streptococcus dysgalactiae, Streptococcus agalactiae,
 CC Streptococcus parabratis, Streptococcus uberis and Streptococcus iniae
 CC (see AAM50665-69). A claimed example is PolyGap4 (see AAM50664).
 CC Expression vectors and host cells for production of the multiple epitope
 CC fusion proteins are provided. The multiple epitope proteins are used in
 CC claimed vaccines for treating or preventing a bacterial infection in a
 CC vertebrate, especially a streptococcal infection, and particularly
 CC mastitis. They are also used in claimed methods of detecting
 CC Streptococcus antibodies. The multiple epitope proteins are capable of
 CC eliciting broad immunity against a variety of streptococcal infections
 CC while minimizing the number of antigens present in the final formulation
 CC and concomitantly reducing production costs
 CC
 XX
 XX Sequence 336 AA:
 SQ
 Query Match 94.3%; Score 1618; DB 5; Length 336;
 Best Local Similarity 92.9%; Pred. No. 4,1e-136;
 Matches 312; Conservative 15; Mismatches 9; Indels 0; Gaps 0;
 QY 1 M V V K G I N G F G R I G R L A F R R I O N V E G V E T R I N D I D P M L A H L K Y D T T O G R F D G T V E V 60
 DB 1 M V V K G I N G F G R I G R L A F R R I O N V E G V E T R I N D I D P M L A H L K Y D T T O G R F D G T V E V 60
 QY 61 K E G G E V N G N F I K V S A R D P E N I D W A T D G V E I V L A L E G T V E V K D G F D W N G K F I V S A E 120
 DB 61 K E G G E V N G N F I K V S A R D P E N I D W A T D G V E I V L A L E G T V E V K D G F D W N G K F I V S A E 120
 QY 121 T A P G A N D V K V V E N T H D I I D G T E V I S A S C T T N C I A P M A K A L H D A F G I O K G L M T T I H A 180

Db 121 TAPGGDDYKTVFNTNHHLLDGETETVLSGASCTTNCCLAPMAKALQDNFVKGQGLMTTTHA 180
 QY 181 YTGDDMTLDGPHRGDGLRRARAGANIVPNSGAAKAGLNTPELNGKLDGAAGRPVPT 240
 Db 181 YTGDDMTLDGPHRGDGLRRARAGANIVPNSGAAKAGLNTPELNGKLDGAAGRPVPT 240
 QY 241 GSVELVYVTLDRNYSVDEINAAKAAANDSGYEDPIVSSDIYVSGSLFPAOTQKVM 300
 Db 241 GSVELVAVLEKETSVSEINAAKAAANDSYGTEDPIVSSDIIGMAYSLFPAOTQKVM 300
 QY 301 EVDGSQLVKKVSVYDNEMSYTAQLVRLTEYFAKIAK 336
 Db 301 TVDGNQLVKKVSVYDNEMSYTAQLVRLTEYFAKIAK 336

RESULT 10

AAMS0641
 ID AAMS0641 standard; protein; 336 AA.

AC AAMS0641;
 XX

DT 04-APR-2002 (first entry)

DE Streptococcus uberis gapC plasmin binding protein.

KM gapC; plasmin-binding protein; UbergapC; infection; mastitis; vaccine;
 KW diagnosis; therapy.

OS Streptococcus uberis.

PN WO200196381-A2.

PD 20-DEC-2001.

PF 11-JUN-2001; 2001WO-CA000838.

PR 12-JUN-2000; 2000US-0211022P.

EA (USA-) UNIV SASKATCHEWAN.

FI Bolton AJ, Perez-Casal J, Fontaine M, Potter AA;
 XX WPI; 2002-130725/17.

DR N-PSDB; ABA91250.

PT Novel isolated gapC protein of Streptococcus dysgalactiae, S. agalactiae,
 ET S. uberis, S. parauberis, or S. iniae, useful as vaccine component for
 FT treating streptococcal infection which causes mastitis in vertebrates.

PS Claim 1(c); Fig 3A-B; 107pp; English.

XX The present sequence is that of the gapC plasmin binding protein

CC (UbergapC) of Streptococcus uberis ATCC 9927, encoded by the gapC gene

CC given in ABA91250. GapC, which has no signal sequence or membrane anchor

CC domain, is capable of eliciting an immune response in a vertebrate. The

CC invention provides the gapC genes and proteins of 5 Streptococcus

CC species, as well as recombinant vectors, host cells and vaccine

CC compositions comprising gapC polynucleotides or proteins. The vaccines

CC are used to treat or prevent a bacterial infection, especially a

CC streptococcal infection, and mastitis in particular (Claimed). GapC

CC proteins are also used in claimed methods for detecting gapC antibodies,

CC and to raise antibodies that are used in claimed methods for detecting

CC gapC proteins. S. uberis is a common pathogen associated with mastitis in

CC cattle, horse, sheep and goat
 XX Sequence 336 AA;
 SQ
 Query Match 94.3%, Score 1618; DB 5; Length 336;
 Best Local Similarity 92.9%; Pred. No. 4.1e-136;
 Matches 312; Conservative 15; Mismatches 9; Indels 0; Gaps 0;
 QY 1 MVAKVGINGEGRIGLAERRIQNVGEVEYRINDLTPNMLHLKXDTQGRFPGTVEV 60

Db 1 MVAKVGINGEGRIGLAERRIQNVGEVEYRINDLTPNMLHLKXDTQGRFPGTVEV 60
 QY 61 KEGFEVNGNFIKVAABDPENDIMATDVEIVLFAATGFAKKEAAEKHLAHANGAKVVI 120
 Db 61 KEGFEVNGNFIKVAABDPENDIMATDVEIVLFAATGFAKKEAAEKHLAHANGAKVVI 120
 QY 121 TAPGANDVKVVFNTNHHLLDGETETVLSGASCTTNCCLAPMAKALQDNFVKGQGLMTTTHA 180
 Db 121 TAPGGDDYKTVFNTNHHLLDGETETVLSGASCTTNCCLAPMAKALQDNFVKGQGLMTTTHA 180
 QY 181 YTGDDMTLDGPHRGDGLRRARAGANIVPNSGAAKAGLNTPELNGKLDGAAGRPVPT 240
 Db 181 YTGDDMTLDGPHRGDGLRRARAGANIVPNSGAAKAGLNTPELNGKLDGAAGRPVPT 240
 QY 241 GSVELVYVTLDRNYSVDEINAAKAAANDSGYEDPIVSSDIYVSGSLFPAOTQKVM 300
 Db 241 GSVELVAVLEKETSVSEINAAKAAANDSYGTEDPIVSSDIIGMAYSLFPAOTQKVM 300
 QY 301 EVDGSQLVKKVSVYDNEMSYTAQLVRLTEYFAKIAK 336
 Db 301 TVDGNQLVKKVSVYDNEMSYTAQLVRLTEYFAKIAK 336

RESULT 11

ABP30758
 ID ABP30758 standard; protein; 336 AA.

AC ABP30758;
 XX

DT 02-JUL-2002 (first entry)

DE Streptococcus polytypide SEQ ID NO 10692.

KM Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
 KW group A streptococcus; Streptococcus pyogenes; antibacterial;
 RM antiinflammatory; infection; vaccine; meningitis; gene therapy.

OS Streptococcus agalactiae.

PN WO200234771-A2.

PD 02-MAY-2002.

PF 29-OCT-2001; 2001WO-GB004789.

PR 27-OCT-2000; 2000GB-00026333.

PR 24-NOV-2000; 2000GB-00028727.

PR 07-MAR-2001; 2001GB-00005640.

PA (CHIR-) CHIRON SPA.
 PA (GENO-) INST GENOMIC RES.

PI Telford J, Masigiani V, Margarit Y Rosl, Grandi G, Fraser C;
 PI Telford J;
 XX

DR WPI; 2002-352536/38.

DR N-PSDB; ABA71389.

XX New Streptococcus protein for the treatment or prevention of infection or

PT disease caused by Streptococcus bacteria, such as meningitis, and for

FT detecting a compound that binds to the protein.

PS Claim 1; Page 4179; 4525pp; English.

CC The invention relates to a protein (ABP25413-ABP30895) from group B

CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS

CC (Streptococcus pyogenes), comprising one of 5483 sequences (SI), given in

CC the specification. The proteins have antibacterial and antiinflammatory

CC activity. (I), nucleic acids encoding (I), ABA66044-ABA71526 and

CC antibodies that bind (I) are used in the manufacture of medicaments for

CC the treatment or prevention of infection or disease caused by

CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.

CC Nucleic acids encoding (I) are used to detect Streptococcus in a

CC biological sample. (1) is used to determine whether a compound binds to
CC (1). A composition comprising (1) or a nucleic acid encoding (1), may be
CC used as a vaccine or diagnostic composition. The disease caused by
CC Streptococcus that is prevented or treated may be meningitis. Nucleic
CC acid encoding (1) may be used to recombinantly produce (1) and may be
CC used in gene therapy. Antibodies to (1) are used for affinity
CC chromatography, immunoassays, and distinguishing/identifying
CC Streptococcus proteins
XX
SQ Sequence 336 AA;

Query Match 93.8%; Score 1609; DB 5; Length 336;
Best Local Similarity 92.0%; Pred. No. 2,6e-135;
Matches 309; Conservative 17; Mismatches 10; Indels 0; Gaps 0;

QY 1 M V V K G I N G F G R I G L A P R I O N V G E V T R I N D L D P N M L A H L K Y D T O G R P D G T V E V 60
DB 1 M V V K G I N G F G R I O L A P R I O N V G E V T R I N D L D P N M L A H L K Y D T O G R P D G T V E V 60
QY 61 K E G G F E V N G F I K V S A E R P E N I D M A T D G V E I V L A T G F P A K E A E K H L A N G A K V I 120
DB 61 K E G G F E V N G F K V S A E R P E N I D M A T D G V E I V L A T G F P A K E A E K H L A N G A K V I 120
QY 121 T A P G N D V K V T V F N T H D I L D G T E V I S G A S C T T N C L A P A K A L Q D N F G V K Q G L M T T I H A 180
DB 121 T A P G N D V K V T V F N T H D I L D G T E V I S G A S C T T N C L A P A K A L Q D N F G V K Q G L M T T I H A 180
QY 181 Y T G D M I L D G P H R G D L R A R A G A N I V P N S T G A A K A I G V I P E L N K L D G A A Q R V P P T 240
DB 181 Y T G D M I L D G P H R G D L R A R A G A N I V P N S T G A A K A I G V I P E L N K L D G A A Q R V P P T 240
QY 241 G S V T E L V T L D K N V S V D E I N A M K A A S N D S F G Y T E D P I V S S D I V G S Y G S L F D A T Q T K V M 300
DB 241 G S V T E L V A T E K V T V E V N A M K A A A N D S Y G T E D P I V S S D I V G S Y G S L F D A T Q T K V M 300
QY 301 E V D S Q L V K V S W Y D N E M S Y T A Q L V T L E Y F A K I A K 336
DB 301 T V D G N Q L V K V S W Y D N E M S Y T A Q L V T L E Y F A K I A K 336

RESULT 12

AA050666
ID AA050666 standard; protein; 336 AA.

AC AA050666;

DT 08-APR-2002 (first entry)

DB Streptococcus agalactiae gapC plasmid binding protein AgalGapC.

XX AgalGapC; GapC; plasmid binding protein; epitope; infection; vaccine;
KW immunisation; mastitis; therapy.

OS Streptococcus agalactiae.

XX WO200196379-A2.

XX 20-DEC-2001.

XX 11-JUN-2001; 2001WO-CA000836.

XX 12-JUN-2000; 2000US-0211247P.

XX (USA-) UNIV SASKATCHEWAN.

XX Potter AA, Perez-Casal J, Fontaine M;

XX WPI; 2002-098053/13.

XX DR N-PSDB; ABA91329.

PT Novel GapC multiple epitope fusion polypeptide comprising antigenic
ET determinant of Streptococcus dysgalactiae, S. agalactiae, S. uberis, S.
ET parameters, or S. iniae GapC protein useful for treating mastitis in

PT vertebrates.

XX Claim 8; Fig 2A-B; 116pp; English.

CC The present sequence is that of the GapC plasmid binding protein.
CC AgalGapC, of Streptococcus agalactiae ATCC 27941. The invention relates
CC to novel GapC multiple epitope fusion proteins that comprise epitopes
CC from 1 or more of Streptococcus dysgalactiae, Streptococcus agalactiae,
CC Streptococcus parvulus, Streptococcus uberis and Streptococcus iniae
CC (see AA050665-69). A claimed example is PolyGap4 (see AA050664).
CC Expression vectors and host cells for production of the multiple epitope
CC fusion proteins are provided. The multiple epitope proteins are used in
CC claimed vaccines for treating or preventing a bacterial infection in a
CC vertebrate, especially a streptococcal infection, and particularly
CC mastitis. They are also used in claimed methods of detecting
CC Streptococcus antibodies. The multiple epitope proteins are capable of
CC eliciting broad immunity against a variety of streptococcal infections
CC while minimizing the number of antigens present in the final formulation
CC and concomitantly reducing production costs

XX Sequence 336 AA;

Query Match 93.4%; Score 1602; DB 5; Length 336;
Best Local Similarity 91.7%; Pred. No. 1.1e-134;
Matches 308; Conservative 17; Mismatches 11; Indels 0; Gaps 0;

QY 1 M V V K G I N G F G R I G L A P R I O N V G E V T R I N D L D P N M L A H L K Y D T O G R P D G T V E V 60
DB 1 M V V V G I N G F G R I G L A P R I O N V G E V T R I N D L D P N M L A H L K Y D T O G R P D G T V E V 60
QY 61 K E G G F E V N G F I K V S A E R P E N I D M A T D G V E I V L A T G F P A K E A E K H L A N G A K V I 120
DB 61 K E G G F E V N Q F K V S A E R P E N I D M A T D G V E I V L A T G F P A K E A E K H L A N G A K V I 120
QY 121 T A P G N D V K V T V F N T H D I L D G T E V I S G A S C T T N C L A P A K A L Q D N F G V K Q G L M T T I H A 180
DB 121 T A P G N D V K V T V F N T H D I L D G T E V I S G A S C T T N C L A P A K A L Q D N F G V K Q G L M T T I H A 180
QY 181 Y T G D M I L D G P H R G D L R A R A G A N I V P N S T G A A K A I G V I P E L N K L D G A A Q R V P P T 240
DB 181 Y T G D M I L D G P H R G D L R A R A G A N I V P N S T G A A K A I G V I P E L N K L D G A A Q R V P P T 240
QY 241 G S V T E L V T L D K N V S V D E I N A M K A A S N D S F G Y T E D P I V S S D I V G S Y G S L F D A T Q T K V M 300
DB 241 G S V T E L V A T E K V T V E V N A M K A A A N D S Y G T E D P I V S S D I V G S Y G S L F D A T Q T K V M 300
QY 301 E V D S Q L V K V S W Y D N E M S Y T A Q L V T L E Y F A K I A K 336
DB 301 T V D G N Q L V K V S W Y D N E M S Y T A Q L V T L E Y F A K I A K 336

RESULT 13

AA050640
ID AA050640 standard; protein; 336 AA.

AC AA050640;

DT 04-APR-2002 (first entry)

DB Streptococcus agalactiae gapC plasmid binding protein.

XX GapC; plasmid-binding protein; AgalGapC; infection; mastitis; vaccine;
KW diagnosis; therapy.

OS Streptococcus agalactiae.

XX WO200196381-A2.

XX 20-DEC-2001.

XX 11-JUN-2001; 2001WO-CA000838.

XX PR 12-JUN-2000; 2000US-0211022P.

XX (UTSA-) UNIV SASKATCHEWAN.
PA Bolton AJ, Perez-Casal J, Fontaine M, Potter AA;
PI
XX
XX WPI: 2002-130725/17.
DR N-PSDB; ABX91249.
XX
XX Novel isolated GapC protein of Streptococcus dysgalactiae, S. agalactiae,
PT S. uberis, S. parauberis, or S. infantiae, useful as vaccine component for
PT treating streptococcal infection which causes mastitis in vertebrates.
XX
PS Claim 1 (b); Fig 2A-B; 107pp; English.
XX
XX The present sequence is that of the GapC plasmid binding protein
CC (AgalGapC) of Streptococcus agalactiae ATCC 27541, encoded by the GapC
CC gene given in ABX91249. GapC, which has no signal sequence or membrane
CC anchor domain, is capable of eliciting an immune response in a
CC vertebrate. The invention provides the GapC genes and proteins of 5
CC Streptococcus species, as well as recombinant vectors, host cells and
CC vaccine compositions comprising GapC polynucleotides or proteins. The
CC vaccines are used to treat or prevent a bacterial infection, especially a
CC streptococcal infection, and mastitis in particular (claimed). GapC
CC proteins are also used in claimed methods for detecting GapC antibodies,
CC and to raise antibodies that are used in claimed methods for detecting
CC GapC proteins. S. agalactiae is a common pathogen associated with
CC mastitis in cattle, horse, sheep and goat. It also causes septicemia,
CC meningitis, bacteraemia, impetigo, arthritis, urinary tract infections,
CC abscesses, spontaneous abortion, etc
XX
SQ Sequence 336 AA:
Query Match 93.4%; Score 1602; DB 5; Length 336;
Best Local Similarity 91.7%; Pred. No. 1,1e-134;
Matches 308; Conservative 17; Mismatches 11; Indels 0; Gaps 0;
QY 1 MYVVGINGRIGRLAERDPEINIDMTDGEIVLEATGFPAKKEAEKRLHANGAKKVI 60
DB 1 MYVVGINGRIGRLAERDPEINIDMTDGEIVLEATGFPAKKEAEKRLHANGAKKVI 60
QY 61 KEGGFVNGKFIKVAERDPEINIDMTDGEIVLEATGFPAKKEAEKRLHANGAKKVI 120
DB 61 KEGGFVNGKFIKVAERDPEINIDMTDGEIVLEATGFPAKKEAEKRLHANGAKKVI 120
QY 121 TAPGNDVKTIVFNTHDLDSTETVLSGASCTTNCIAPAKAHLDAFGIQLGTLTTTHA 180
DB 121 TAPGNDVKTIVFNTHDLDSTETVLSGASCTTNCIAPAKAHLDAFGIQLGTLTTTHA 180
QY 181 YTGDMTLDGPHRGSPLRRARAGANIVPNSGAAKAGLVIPELNKLDAAGORVPVPT 240
DB 181 YTGDMTLDGPHRGSPLRRARAGANIVPNSGAAKAGLVIPELNKLDAAGORVPVPT 240
QY 241 GSVELVATLEKQVTEEVNAAKAAANDSGVTEDEPIVSDIVGISGLFPAQTQVQ 300
DB 241 GSVELVATLEKQVTEEVNAAKAAANDSGVTEDEPIVSDIVGISGLFPAQTQVQ 300
QY 301 EVDGSLIVKVSIVTDEKSYTQVRLTEPAKIAK 336
DB 301 EVDGSLIVKVSIVTDEKSYTQVRLTEPAKIAK 336
QY 301 TVDGNQIVKVSIVTDEKSYTQVRLTEPAKIAK 336
DB 301 TVDGNQIVKVSIVTDEKSYTQVRLTEPAKIAK 336
RESULT 14
AB002516
ID AB002516 standard; protein; 335 AA.
XX
XX AB002516;
XX
XX 23-OCT-2003 (revised)
DT 11-FEB-2003 (first entry)
XX
XX S. pneumoniae type 4 strain protein from coding region #2094.
DE
XX
XX Bacterial meningitis; pneumonia; sepsis; otitis media; ear infection;
KV

KM antiinflammatory; antibacterial; immunostimulant; auditory; respiratory;
KM gene therapy; vaccine.
XX
XX Streptococcus pneumoniae; type 4 strain.
OS
XX
XX WO200277021-A2.
PM
XX
XX 03-OCT-2002.
PD
XX
XX 27-MAR-2002; 2002MO-IB002163.
PF
XX
XX 27-MAR-2001; 2001GB-00007658.
PR
XX
XX (CHIR-) CHIRON SPA.
PA (GENO-) INST GENOMIC RES.
PI Masignani V, Tettelin H, Fraser C;
XX
XX WPI: 2003-040579/03.
DR N-PSDB; ABX07806.
XX
XX New proteins and nucleic acid molecules from Streptococcus pneumoniae,
PT useful as medicaments for treating or preventing a disease or infection
PT due to Streptococcus bacteria, such as pneumonia, sepsis, otitis media or
PT ear infection.
XX
PS Claim 1; SEQ ID NO 4189; 56pp; English.
XX
XX The invention relates to a protein comprising or having at least 50%
CC identity to any of the 2469 amino acid sequences, identified in the
CC specification (available on a computer readable format), or its fragment,
CC expressed from 2469 of 2489 identified DNA coding regions from the
CC Streptococcus pneumoniae type 4 strain genomic sequence appearing as
CC ABX58454. Also included are an antibody which binds one of the proteins,
CC treating a patient by administering the protein, DNA or antibody (in a
CC composition), a kit comprising first and second primers, which are the
CC nucleic acid cited above or fragments between nucleotides 8-100 of a
CC sequence not defined in the specification, for amplifying a target
CC sequence contained within a Streptococcus nucleic acid sequence, where
CC the first primer is substantially complementary to the target sequence
CC and the second primer is substantially complementary to the complement of
CC the target sequence, and where the parts of the primers having
CC substantial complementarity define the termini of the target sequence to
CC be amplified, assay comprising contacting a test compound with the
CC protein, and determining whether the test compound binds to the protein
CC and a Streptococcus pneumoniae bacterium, where one or more genes
CC encoding the protein has been rendered inactive. The proteins, nucleic
CC acid molecules, antibody and compositions are useful as medicaments for
CC treating or preventing a disease or infection due to Streptococcus
CC bacteria, particularly S. pneumoniae, such as pneumonia, sepsis, otitis
CC media or ear infection. They are also useful in developing vaccines,
CC diagnostics and antibiotics. The methods are useful for identifying
CC immunodominant proteins. The present sequence is one of the 2469 proteins
CC expressed by the identified coding regions from the genomic sequence.
CC Note: The sequence data for this patent did not form part of the printed
CC at fip.wipo.int/publ/published_pct_sequences. (Updated on 23-OCT-2003 to
CC standardise OS field)
XX
SQ Sequence 335 AA:
Query Match 92.9%; Score 1593.5; DB 6; Length 335;
Best Local Similarity 92.6%; Pred. No. 6.3e-134;
Matches 311; Conservative 12; Mismatches 12; Indels 1; Gaps 1;
QY 1 MYVVGINGRIGRLAERDPEINIDMTDGEIVLEATGFPAKKEAEKRLHANGAKKVI 60
DB 1 MYVVGINGRIGRLAERDPEINIDMTDGEIVLEATGFPAKKEAEKRLHANGAKKVI 60
QY 61 KEGGFVNGKFIKVAERDPEINIDMTDGEIVLEATGFPAKKEAEKRLHANGAKKVI 120
DB 61 KEGGFVNGKFIKVAERDPEINIDMTDGEIVLEATGFPAKKEAEKRLHANGAKKVI 120

QY 121 TAPGNDVKTIVFNTNHDILDTGETVYISGASCTTNCIAPAKAALHDAFGIOKGMTTTHA 180
 DB 120 TAPGNDVKTIVFNTNHDILDTGETVYISGASCTTNCIAPAKAALHDAFGIOKGMTTTHA 179
 QY 181 YTGQMLIDGPHRGDLPRAAGANIVPNSGAAKALGIVPELNGKLDGSAORVPEPT 240
 DB 180 YTGQMLIDGPHRGDLPRAAGANIVPNSGAAKALGIVPELNGKLDGSAORVPEPT 239
 QY 241 GSVELVAVTLDKRVSYDEINAMKASNSFGYTEDPIVSSDIYGVSYSLPDATQTKVM 300
 DB 240 GSVELVAVTLDKRVSYDEINAMKASNSFGYTEDPIVSSDIYGVSYSLPDATQTKVM 299
 QY 301 EVDGSOLVKVSWYDNEMSYTAQLVRLTEYFAKIAK 336
 DB 300 DVDGQLVKVVSWYDNEMSYTAQLVRLTEYFAKIAK 335

RESULT 15
 AAU37576
 ID AAU37576 standard; protein: 359 AA.
 XX
 AC AAU37576;
 XX
 DT 14-FEB-2002 (first entry)
 XX
 DE Streptococcus pneumoniae cellular proliferation protein #5.
 XX
 KW Antisense; prokaryotic cellular proliferation protein; antibiotic;
 XX antibacterial; drug design.
 XX
 OS Streptococcus pneumoniae.
 XX
 PN MO200170355-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 21-MAR-2001; 2001MO-US009180.
 XX
 PR 21-MAR-2000; 2000US-0191078P.
 XX 23-MAY-2000; 2000US-0206848P.
 PR 26-MAY-2000; 2000US-0207727P.
 PR 23-OCT-2000; 2000US-0242578P.
 PR 27-NOV-2000; 2000US-0253625P.
 PR 22-DEC-2000; 2000US-0257931P.
 PR 16-FEB-2001; 2001US-0269308P.
 XX
 EA (BLIT-) ELITRA PHARM INC.
 XX
 PI Haselbeck R, Ohlsen KL, Zyskind TW, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;
 XX
 DR MPI; 2001-611495/70.
 XX
 DR N-PSDB; AAS55435.
 XX
 PT New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids.
 XX
 PS Example 3; SEQ ID NO 13169; S11PP; English.
 XX
 CC The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the genes,
 CC their use in the discovery of novel antibiotics, the essential genes
 CC themselves and the encoded proteins. The prokaryotes used are Escherichia
 CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
 CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
 CC useful for the identification of potential new targets for antibiotic
 CC development. The antisense nucleic acids can also be used to identify
 CC proteins used in proliferation, to express these proteins, and to obtain
 CC antibodies capable of binding to the expressed proteins. The proteins can
 CC be used to screen compounds in rational drug discovery programmes. The
 CC antisense nucleic acid sequence is also useful to screen for homologous
 CC nucleic acids which are required for cell proliferation in a wide variety
 CC of organisms. The present sequence represents an essential prokaryotic

CC cellular proliferation protein. Note: The sequence data for this patent
 CC did not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp://wipo.int/pub/published_pct_sequences

XX
 SQ Sequence 359 AA;
 XX
 Query Match 92.9%; Score 1593.5; DB 4; Length 359;
 Best Local Similarity 92.6%; Pred. No. 6.9e-134;
 Matches 311; Conservative 12; Mismatches 12; Indels 1; Gaps 1;

QY 1 MYKVGINGPGRIGRLAFRIQVGVGVTRINDLTPMNLAKLYDTTQGRFDGTEV 60
 DB 25 MYKVGINGPGRIGRLAFRIQVGVGVTRINDLTPMNLAKLYDTTQGRFDGTEV 84
 QY 61 KEGFEVNGNFTKVSARDPENIDATGVEIVLEATGFAKKEAAEGLHANKAKVVI 120
 DB 85 KEGFEVNGNFTKVSARDPENIDATGVEIVLEATGFAKKEAAEGLHANKAKVVI 143
 QY 121 TAPGNDVKTIVFNTNHDILDTGETVYISGASCTTNCIAPAKAALHDAFGIOKGMTTTHA 180
 DB 144 TAPGNDVKTIVFNTNHDILDTGETVYISGASCTTNCIAPAKAALHDAFGIOKGMTTTHA 203
 QY 181 YTGQMLIDGPHRGDLPRAAGANIVPNSGAAKALGIVPELNGKLDGSAORVPEPT 240
 DB 204 YTGQMLIDGPHRGDLPRAAGANIVPNSGAAKALGIVPELNGKLDGSAORVPEPT 263
 QY 241 GSVELVAVTLDKRVSYDEINAMKASNSFGYTEDPIVSSDIYGVSYSLPDATQTKVM 300
 DB 264 GSVELVAVTLDKRVSYDEINAMKASNSFGYTEDPIVSSDIYGVSYSLPDATQTKVM 323
 QY 301 EVDGSOLVKVSWYDNEMSYTAQLVRLTEYFAKIAK 336
 DB 324 DVDGQLVKVVSWYDNEMSYTAQLVRLTEYFAKIAK 359

Search completed: May 5, 2004, 14:19:41
 Job time : 60 secs

GenCore Version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 14:18:39 ; Search time 23 Seconds

(Without alignments)
754.188 Million cell updates/sec

Title: US-09-878-781-4

Perfect score: 1715

Sequence: 1 MVYKVGINGGRIGRLAFRR.....EMSYTAQLVRLTFEFAXIAK 336

Scoring table:

BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database:

Issued Patents AA:*
1: /cgm2_6/ptodata/2/1aa/5A COMB pep:*
2: /cgm2_6/ptodata/2/1aa/5B COMB pep:*
3: /cgm2_6/ptodata/2/1aa/6A COMB pep:*
4: /cgm2_6/ptodata/2/1aa/6B COMB pep:*
5: /cgm2_6/ptodata/2/1aa/6C COMB pep:*
6: /cgm2_6/ptodata/2/1aa/6D COMB pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1715	100.0	336	4	US-09-878-766A-12
2	1711	99.8	336	1	US-07-928-462-2
3	1711	99.8	336	3	US-08-273-247-2
4	1656.5	96.6	448	4	US-09-878-766A-22
5	1618	94.3	336	4	US-09-878-766A-16
6	1602	93.4	336	4	US-09-878-766A-14
7	1584	92.4	336	4	US-09-878-766A-20
8	1578	92.0	336	4	US-09-878-766A-18
9	1549.5	90.3	333	3	US-08-961-083-54
10	1549.5	90.3	333	4	US-09-536-784-54
11	1337.5	78.0	346	4	US-09-134-000C-4400
12	1335.5	77.3	333	4	US-09-107-532A-4769
13	1190	69.4	348	4	US-09-134-001C-5513
14	967	56.4	357	4	US-09-134-000C-4229
15	966	56.3	340	4	US-09-634-238-233
16	959	55.9	349	4	US-09-489-039A-13939
17	896.5	52.3	335	6	5290690-9
18	885.5	51.6	334	6	5290690-10
19	835.5	48.7	334	6	5290690-11
20	830	48.4	335	4	US-09-489-039A-7679
21	824.5	48.1	340	3	US-09-095-855-205
22	824.5	48.1	340	4	US-09-205-426-205
23	799	46.6	340	4	US-09-543-681A-8083
24	770.5	44.9	352	4	US-09-489-039A-9410
25	766.5	44.7	282	4	US-09-107-532A-4057
26	740.5	43.2	336	4	US-09-198-452A-662
27	736	42.9	338	4	US-09-091-725-10

28	718.5	41.9	335	2	US-08-903-800A-6	Sequence 6, Appl
29	715	41.7	340	4	US-09-252-991A-21612	Sequence 21612, A
30	709.5	41.4	334	4	US-09-674-826B-6	Sequence 6, Appl
31	699.5	40.8	268	2	US-08-997-080-186	Sequence 186, App
32	699.5	40.8	268	2	US-08-997-362-186	Sequence 186, App
33	699.5	40.8	268	3	US-09-095-855-186	Sequence 186, App
34	699.5	40.8	268	4	US-09-324-542-186	Sequence 186, App
35	699.5	40.8	268	4	US-09-205-426-186	Sequence 186, App
36	697.5	40.7	341	4	US-09-134-001C-3783	Sequence 11, Appl
37	667.5	38.9	333	3	US-09-532-803-9	Sequence 9, Appl
38	659.5	38.5	333	3	US-09-548-938A-13	Sequence 13, Appl
39	645	37.6	277	4	US-09-489-039A-11854	Sequence 11854, A
40	580	33.8	350	4	US-09-328-352-6704	Sequence 6704, Ap
41	572.5	33.4	340	4	US-09-540-236-3378	Sequence 3378, Ap
42	570	33.2	442	4	US-09-540-236-3378	Sequence 3470, Ap
43	553	32.2	417	4	US-09-543-681A-6082	Sequence 6082, Ap
44	552	32.2	359	4	US-09-252-991A-19846	Sequence 19846, A
45	549	32.0	367	4	US-09-252-991A-19846	Sequence 19846, A

ALIGNMENTS

RESULT 1
US-09-878-766A-12
; Sequence 12, Application US/09878766A
; Patent No. 6660270
; GENERAL INFORMATION:
; APPLICANT: Pottery, Andrew A.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GABC PROTEIN
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/09/878, 766A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 12
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus dysgalactiae
US-09-878-766A-12

Query Match 100.0%; Score 1715; DB 4; Length 336;
Best Local Similarity 100.0%; Pred. No. 4.2e-164;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVYKVGINGGRIGRLAFRRIONVEYTRINDLTDPMALHLKDTTQRPFGTYEV 60
DB 1 MVYKVGINGGRIGRLAFRRIONVEYTRINDLTDPMALHLKDTTQRPFGTYEV 60
QY 61 KEGGPNVSPITVSERDPENIDMTDGEIVLEATGEPFAKKEAKKHLHANGAKKVI 120
DB 61 KEGGPNVSPITVSERDPENIDMTDGEIVLEATGEPFAKKEAKKHLHANGAKKVI 120
QY 121 TARGNDVKTIVNTNHDILDTETVLSGASCTTNCALPMKALHDAFGIOKGMTTHA 180
DB 121 TARGNDVKTIVNTNHDILDTETVLSGASCTTNCALPMKALHDAFGIOKGMTTHA 180
QY 181 YTGDMILDFHRRGGLRRARAGAAIVPMSSTAAGVITPELNCKLDGAORVPVT 240
DB 181 YTGDMILDFHRRGGLRRARAGAAIVPMSSTAAGVITPELNCKLDGAORVPVT 240
QY 181 YTGDMILDFHRRGGLRRARAGAAIVPMSSTAAGVITPELNCKLDGAORVPVT 240
DB 181 YTGDMILDFHRRGGLRRARAGAAIVPMSSTAAGVITPELNCKLDGAORVPVT 240
QY 241 GSVEITVYTLDSKVSVDENINAMKAANDSFYTBPDIVSSDIVGVSGLFDATQTXM 300
DB 241 GSVEITVYTLDSKVSVDENINAMKAANDSFYTBPDIVSSDIVGVSGLFDATQTXM 300
QY 301 EVDGSQIVKVSYSYDENSYSYTAQLVRLTFEFAXIAK 336
DB 301 EVDGSQIVKVSYSYDENSYSYTAQLVRLTFEFAXIAK 336

RESULT 2

US-07-928-462-2
Sequence 2, Application US/07928462
Patent No. 5328996
GENERAL INFORMATION:
APPLICANT: Boyle, Michael D.P.
APPLICANT: Lottenberg, Richard
APPLICANT: Broder, Christopher C.
APPLICANT: von Mering, Gregory O.
TITLE OF INVENTION: Bacterial Plasmid Receptors as
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/928,462
FILING DATE: 19920810
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/524,411
FILING DATE: 16-MAY-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/330,849
FILING DATE: 29-MAR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: UF/S&S-13.C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 336 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-928-462-2

Query Match 99.8%; Score 1711; DB 1; Length 336;
Best Local Similarity 99.4%; Pred. No. 1,le-163;
Matches 334; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYVKGINGRGRIGR.AFRRIQNVGVEVTRINDLTDPMALHLKDTTQGRFGYEV 60
DB 1 MYVKGINGRGRIGR.AFRRIQNVGVEVTRINDLTDPMALHLKDTTQGRFGYEV 60
QY 61 KEGGFVNGNFIKVSARDPENIDATDGVETLVEATGFPKKEAEKHLHANGAKKVI 120
DB 61 KEGGFVNGNFIKVSARDPENIDATDGVETLVEATGFPKKEAEKHLHANGAKKVI 120
QY 121 TAPGNDVKTIVNINHDILDTETVLSGASCTTNCLAPMAKALHDAFGIOKGLMTTTHA 180
DB 121 TAPGNDVKTIVNINHDILDTETVLSGASCTTNCLAPMAKALHDAFGIOKGLMTTTHA 180
QY 181 YTGDMILDGHRGGLRRARAGANTVPSNTAAKALIGLVIPENKLDGAAGRVPPT 240
DB 181 YTGDMILDGHRGGLRRARAGANTVPSNTAAKALIGLVIPENKLDGAAGRVPPT 240
QY 241 GSVTELVVTLDKVSVDELNSMKASNDSTFYTHLPVSSDIGVSGSLPDAQTQVM 300
DB 241 GSVTELVVTLDKVSVDELNSMKASNDSTFYTHLPVSSDIGVSGSLPDAQTQVM 300

QY 301 EVDGSQLVYVSWYDNMSYTAQVLTLEYPAKTAX 336
DB 301 EVDGSQLVYVSWYDNMSYTAQVLTLEYPAKTAX 336

RESULT 3

US-08-273-247-2
Sequence 2, Application US/08273247
Patent No. 6136323
GENERAL INFORMATION:
APPLICANT: Boyle, Michael D.P.
APPLICANT: Lottenberg, Richard
APPLICANT: Broder, Christopher C.
APPLICANT: von Mering, Gregory O.
TITLE OF INVENTION: Bacterial Plasmid Receptors as
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/273,247
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/524,411
FILING DATE: 16-MAY-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/330,849
FILING DATE: 29-MAR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: UF/S&S-13.C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 336 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-273-247-2

Query Match 99.8%; Score 1711; DB 3; Length 336;
Best Local Similarity 99.4%; Pred. No. 1,le-163;
Matches 334; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYVKGINGRGRIGR.AFRRIQNVGVEVTRINDLTDPMALHLKDTTQGRFGYEV 60
DB 1 MYVKGINGRGRIGR.AFRRIQNVGVEVTRINDLTDPMALHLKDTTQGRFGYEV 60
QY 61 KEGGFVNGNFIKVSARDPENIDATDGVETLVEATGFPKKEAEKHLHANGAKKVI 120
DB 61 KEGGFVNGNFIKVSARDPENIDATDGVETLVEATGFPKKEAEKHLHANGAKKVI 120
QY 121 TAPGNDVKTIVNINHDILDTETVLSGASCTTNCLAPMAKALHDAFGIOKGLMTTTHA 180
DB 121 TAPGNDVKTIVNINHDILDTETVLSGASCTTNCLAPMAKALHDAFGIOKGLMTTTHA 180
QY 181 YTGDMILDGHRGGLRRARAGANTVPSNTAAKALIGLVIPENKLDGAAGRVPPT 240
DB 181 YTGDMILDGHRGGLRRARAGANTVPSNTAAKALIGLVIPENKLDGAAGRVPPT 240

QY 241 GSVTELVTLDRKVSDEINAAKKAASNDSPGYTEDPYISDVGYSGLFPAQTQKM 300
DB 241 GSVTELVTLDRKVSDEINAAKKAASNDSPGYTEDPYISDVGYSGLFPAQTQKM 300
QY 301 EVDGSQLVKVSWMNEMSYTAQVLTLEYPAKIAK 336
DB 301 EVDGSQLVKVSWMNEMSYTAQVLTLEYPAKIAK 336

RESULT 4

US-09-878-766A-22
; Sequence 22, Application US/09878766A
; Patent No. 6660270
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/09/878,766A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 448
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GapC multiple
US-09-878-766A-22

Query Match 96.6%; Score 1656.5; DB 4; Length 448;
Best Local Similarity 79.6%; Pred. No. 5.1e-158;
Matches 335; Conservative 0; Mismatches 1; Indels 85; Gaps 1;

QY 1 MVAKGINSGRIGRLAFRRIONVGEVTRINDLIDPMLAHLKXDTTQGRPGCTVEV 60
DB 23 MVAKGINSGRIGRLAFRRIONVGEVTRINDLIDPMLAHLKXDTTQGRPGCTVEV 87
QY 61 KEGGEVNGFIVKSAERDPE----- 81
DB 88 KEGGEVNGFIVKSAERDPE----- 147
QY 82 -----NIMATGCVETIYIE 95
DB 148 KQPEQIDWATDVEIVLEIDGTVEVEGGEFVNGQFVKVSAEREPANIDWATDVEIVLE 207
QY 96 ANGEFFAKKEAARKEKHLHANGAKKRVITAPGNDVTVVPTNNDILDGTEVIVSAGCTTN 155
DB 208 AISEFKEKKAARKEKHLHANGAKKRVITAPGNDVTVVPTNNDILDGTEVIVSAGCTTN 267
QY 156 CLAPAAKALHDAFGIOKGLMTTTHAYTGDQMLIDGPHRGDLRRARAGANIVPNSGAA 215
DB 268 CLAPAAKALHDAFGIOKGLMTTTHAYTGDQMLIDGPHRGDLRRARAGANIVPNSGAA 327
QY 216 KALGVITPELNGKLDGAAGRVPPVPGSTTEIVTLDRKVSDEINAAKKAASNDSPGYTE 275
DB 328 KALGVITPELNGKLDGAAGRVPPVPGSTTEIVTLDRKVSDEINAAKKAASNDSPGYTE 387
QY 276 DPTVSSDIIVGYSGSLFDATQTKWEVDGSQLVVSWMNEMSYTAQVLTLEYPAKIA 335
DB 388 DPTVSSDIIVGYSGSLFDATQTKWEVDGSQLVVSWMNEMSYTAQVLTLEYPAKIA 447
QY 336 K 336
DB 448 K 448

RESULT 5
US-09-878-766A-16

; Sequence 16, Application US/09878766A
; Patent No. 6660270
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/09/878,766A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus uberis
US-09-878-766A-16

Query Match 94.3%; Score 1618; DB 4; Length 336;
Best Local Similarity 92.9%; Pred. No. 2.4e-154;
Matches 312; Conservative 15; Mismatches 9; Indels 0; Gaps 0;

QY 1 MVAKGINSGRIGRLAFRRIONVGEVTRINDLIDPMLAHLKXDTTQGRPGCTVEV 60
DB 1 MVAKGINSGRIGRLAFRRIONVGEVTRINDLIDPMLAHLKXDTTQGRPGCTVEV 60
QY 61 KEGGEVNGFIVKSAERDPEINIDWATDVEIVLEIDGTVEVEGGEFVNGQFVKVSAEREPANIDWATDVEIVLE 120
DB 61 KEGGEVNGFIVKSAERDPEINIDWATDVEIVLEIDGTVEVEGGEFVNGQFVKVSAEREPANIDWATDVEIVLE 120
QY 121 TAPGNDVKTIVPTNNDILDGTEVIVSAGCTTNCLAPAAKALHDAFGIOKGLMTTTHA 180
DB 121 TAPGNDVKTIVPTNNDILDGTEVIVSAGCTTNCLAPAAKALHDAFGIOKGLMTTTHA 180
QY 181 YTGQMTLIDGPHRGDLRRARAGANIVPNSGAAKALGVITPELNGKLDGAAGRVPPVPT 240
DB 181 YTGQMTLIDGPHRGDLRRARAGANIVPNSGAAKALGVITPELNGKLDGAAGRVPPVPT 240
QY 241 GSVTELVTLDRKVSDEINAAKKAASNDSPGYTEDPYISDVGYSGLFPAQTQKM 300
DB 241 GSVTELVTLDRKVSDEINAAKKAASNDSPGYTEDPYISDVGYSGLFPAQTQKM 300
QY 301 EVDGSQLVKVSWMNEMSYTAQVLTLEYPAKIAK 336
DB 301 EVDGSQLVKVSWMNEMSYTAQVLTLEYPAKIAK 336

RESULT 6

US-09-878-766A-14
; Sequence 14, Application US/09878766A
; Patent No. 6660270
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/09/878,766A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus agalactiae
US-09-878-766A-14

Query Match 93.4%; Score 1602; DB 4; Length 336;
Best Local Similarity 91.7%; Pred. No. 9.9e-153;
Matches 308; Conservative 17; Mismatches 11; Indels 0; Gaps 0;

QY 1 M V V K V G I N G F R I G R I A F R I O N V E G V E T R I N D L T D P M L A H L K Y D T T G R F D G T V E V 60
Db 1 M V V K V G I N G F R I G R I A F R I O N V E G V E T R I N D L T D P M L A H L K Y D T T G R F D G T V E V 60
QY 61 K E G G F E V N G F I K V S A R E P D E N I M A T D G V E I V L E A T G F P A K K A A B H L H N A K K V I 120
Db 61 K E G G F E V N G F I K V S A R E P D E N I M A T D G V E I V L E A T G F P A K K A A B H L H N A K K V I 120
QY 121 T A P G N D V K T V E N T N H D I L D G T E T V I S G A S C T T N C L A P M A K A L D A F G I O K G L M T T I H A 180
Db 121 T A P G N D V K T V E N T N H D I L D G T E T V I S G A S C T T N C L A P M A K A L D A F G I O K G L M T T I H A 180
QY 181 Y T G D M I L D G P H R G D L R R A R A G A N I V P N S T G A A K A I G V I P E L N G K L D G A A G R V P V P T 240
Db 181 Y T G D M I L D G P H R G D L R R A R A G A N I V P N S T G A A K A I G V I P E L N G K L D G A A G R V P V P T 240
QY 241 G S V T E L V A V L E K D T V E E V N A M A K A A N D S Y G T E D P I V S D I V G S L P D A T Q T K V Q 300
Db 241 G S V T E L V A V L E K D T V E E V N A M A K A A N D S Y G T E D P I V S D I V G S L P D A T Q T K V Q 300
QY 301 E V D G S Q L V K V S W T D N E M S Y T A Q L V R T L E Y F A K I A K 336
Db 301 T V D G N Q L V K V S W T D N E M S Y T A Q L V R T L E Y F A K I A K 336

RESULT 7
US-09-878-766A-20
Sequence 20, Application US/09878766A
Patent No. 6660270
GENERAL INFORMATION:
APPLICANT: Potter, Andrew A.
APPLICANT: Perez-Casal, Jose
APPLICANT: Fontaine, Michael
TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION
FILE REFERENCE: 9000-0057
CURRENT APPLICATION NUMBER: US/09/878, 766A
CURRENT FILING DATE: 2001-09-10
NUMBER OF SEQ ID NOS: 22
SOFTWARE: PatentIn ver. 2.0
SEQ ID NO 20
LENGTH: 336
TYPE: PRT
ORGANISM: Streptococcus iniiae
US-09-878-766A-20

Query Match 92.4%; Score 1584; DB 4; Length 336;
Best Local Similarity 90.5%; Pred. No. 6.4e-151;
Matches 304; Conservative 19; Mismatches 13; Indels 0; Gaps 0;
QY 1 M V V K V G I N G F R I G R I A F R I O N V E G V E T R I N D L T D P M L A H L K Y D T T G R F D G T V E V 60
Db 1 M V V K V G I N G F R I G R I A F R I O N V E G V E T R I N D L T D P M L A H L K Y D T T G R F D G T V E V 60
QY 61 K E G G F E V N G F I K V S A R E P D E N I M A T D G V E I V L E A T G F P A K K A A B H L H N A K K V I 120
Db 61 K E G G F E V N G F I K V S A R E P D E N I M A T D G V E I V L E A T G F P A K K A A B H L H N A K K V I 120
QY 121 T A P G N D V K T V E N T N H D I L D G T E T V I S G A S C T T N C L A P M A K A L D A F G I O K G L M T T I H A 180
Db 121 T A P G N D V K T V E N T N H D I L D G T E T V I S G A S C T T N C L A P M A K A L D A F G I O K G L M T T I H A 180
QY 181 Y T G D M I L D G P H R G D L R R A R A G A N I V P N S T G A A K A I G V I P E L N G K L D G A A G R V P V P T 240
Db 181 Y T G D M I L D G P H R G D L R R A R A G A N I V P N S T G A A K A I G V I P E L N G K L D G A A G R V P V P T 240
QY 241 G S V T E L V A V L E K D T V E E V N A M A K A A N D S Y G T E D P I V S D I V G S L P D A T Q T K V Q 300
Db 241 G S V T E L V A V L E K D T V E E V N A M A K A A N D S Y G T E D P I V S D I V G S L P D A T Q T K V Q 300
QY 301 E V D G S Q L V K V S W T D N E M S Y T A Q L V R T L E Y F A K I A K 336
Db 301 T V D G N Q L V K V S W T D N E M S Y T A Q L V R T L E Y F A K I A K 336

RESULT 8
US-09-878-766A-18
Sequence 18, Application US/09878766A
Patent No. 6660270
GENERAL INFORMATION:
APPLICANT: Potter, Andrew A.
APPLICANT: Perez-Casal, Jose
APPLICANT: Fontaine, Michael
TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION
FILE REFERENCE: 9000-0057
CURRENT APPLICATION NUMBER: US/09/878, 766A
CURRENT FILING DATE: 2001-09-10
NUMBER OF SEQ ID NOS: 22
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 18
LENGTH: 336
TYPE: PRT
ORGANISM: Streptococcus parauberis
US-09-878-766A-18

Query Match 92.0%; Score 1578; DB 4; Length 336;
Best Local Similarity 90.8%; Pred. No. 2.6e-150;
Matches 305; Conservative 16; Mismatches 15; Indels 0; Gaps 0;
QY 1 M V V K V G I N G F R I G R I A F R I O N V E G V E T R I N D L T D P M L A H L K Y D T T G R F D G T V E V 60
Db 1 M V V K V G I N G F R I G R I A F R I O N V E G V E T R I N D L T D P M L A H L K Y D T T G R F D G T V E V 60
QY 61 K E G G F E V N G F I K V S A R E P D E N I M A T D G V E I V L E A T G F P A K K A A B H L H N A K K V I 120
Db 61 K E G G F E V N G F I K V S A R E P D E N I M A T D G V E I V L E A T G F P A K K A A B H L H N A K K V I 120
QY 121 T A P G N D V K T V E N T N H D I L D G T E T V I S G A S C T T N C L A P M A K A L D A F G I O K G L M T T I H A 180
Db 121 T A P G N D V K T V E N T N H D I L D G T E T V I S G A S C T T N C L A P M A K A L D A F G I O K G L M T T I H A 180
QY 181 Y T G D M I L D G P H R G D L R R A R A G A N I V P N S T G A A K A I G V I P E L N G K L D G A A G R V P V P T 240
Db 181 Y T G D M I L D G P H R G D L R R A R A G A N I V P N S T G A A K A I G V I P E L N G K L D G A A G R V P V P T 240
QY 241 G S V T E L V A V L E K D T V E E V N A M A K A A N D S Y G T E D P I V S D I V G S L P D A T Q T K V Q 300
Db 241 G S V T E L V A V L E K D T V E E V N A M A K A A N D S Y G T E D P I V S D I V G S L P D A T Q T K V Q 300
QY 301 E V D G S Q L V K V S W T D N E M S Y T A Q L V R T L E Y F A K I A K 336
Db 301 T V D G N Q L V K V S W T D N E M S Y T A Q L V R T L E Y F A K I A K 336

RESULT 9
US-08-961-083-54
Sequence 54, Application US/08961083
Patent No. 6159469
GENERAL INFORMATION:
APPLICANT: Choi et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines
NUMBER OF SEQUENCES: 452
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/961,083
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB34022
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 333 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-961-083-54

Query Match 90.3%; Score 1549.5; DB 3; Length 333;
Best Local Similarity 91.3%; Pred. No. 1.9e-147;
Matches 303; Conservative 12; Mismatches 16; Indels 1; Gaps 1;

QY 2 VVKVINGFGRIQRLAFRIQNVGEVETRIINDLTPNNMLAHLYKYDTTQGRDGTVEVK 61
DB 1 VVKVINGFGRIQRLAFRIQNVGEVETRIINDLTPNNMLAHLYKYDTTQGRDGTVEVK 60
QY 62 EGGFEVNGEIKVSAREPDIWATDGEIVLEATGFFAKKEAEKHLHANGAKKVIT 121
DB 61 EGGFEVNGEIKVSAREPDIWATDGEIVLEATGFFAKKEAEKHLHANGAKKVIT 119
QY 122 AFGNDVKTIVFNTNHDILDTETVTSASCTTNCCLAPMAKALHDAFGIOKGLTTIHAY 181
DB 120 AFGNDVKTIVFNTNHDILDTETVTSASCTTNCCLAPMAKALHDAFGIOKGLTTIHAY 179
QY 182 TSDQWILDGPHRGDRLRAGAGANIVPNSGAAKAGIVTPELNGKLDGAAQRPVPTG 241
DB 180 TSDQWILDGPHRGDRLRAGAGANIVPNSGAAKAGIVTPELNGKLDGAAQRPVPTG 239
QY 242 SYTELAVTLDRKNSVDEINAAKKAASNDGFTYEDPIVSSDIVGVSGSLFPAOTQKWE 301
DB 240 SYTELAVTLDRKNSVDEINAAKKAASNDGFTYEDPIVSSDIVGVSGSLFPAOTQKWE 299
QY 302 VDSGQIVKVSNTNENSTYTAQLVETLEPAK 333
DB 300 VDSGQIVKVSNTNENSTYTAQLVETLEPAK 331

RESULT 10
US-09-536-784-54
Sequence 54, Application US/09556784
Patent No. 6573082

GENERAL INFORMATION:
APPLICANT: Choi et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
NUMBER OF SEQUENCES: 452
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/536,784

FILING DATE: 30-Oct-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/961,083
FILING DATE: OCT-30-1997
ATTORNEY/AGENT INFORMATION:
NAME: Michelle S. Marks
REGISTRATION NUMBER: 41,971
REFERENCE/DOCKET NUMBER: PB340P3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 333 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 54:
US-09-536-784-54

Query Match 90.3%; Score 1549.5; DB 4; Length 333;
Best Local Similarity 91.3%; Pred. No. 1.9e-147;
Matches 303; Conservative 12; Mismatches 16; Indels 1; Gaps 1;

QY 2 VVKVINGFGRIQRLAFRIQNVGEVETRIINDLTPNNMLAHLYKYDTTQGRDGTVEVK 61
DB 1 VVKVINGFGRIQRLAFRIQNVGEVETRIINDLTPNNMLAHLYKYDTTQGRDGTVEVK 60
QY 62 EGGFEVNGEIKVSAREPDIWATDGEIVLEATGFFAKKEAEKHLHANGAKKVIT 121
DB 61 EGGFEVNGEIKVSAREPDIWATDGEIVLEATGFFAKKEAEKHLHANGAKKVIT 119
QY 122 AFGNDVKTIVFNTNHDILDTETVTSASCTTNCCLAPMAKALHDAFGIOKGLTTIHAY 181
DB 120 AFGNDVKTIVFNTNHDILDTETVTSASCTTNCCLAPMAKALHDAFGIOKGLTTIHAY 179
QY 182 TSDQWILDGPHRGDRLRAGAGANIVPNSGAAKAGIVTPELNGKLDGAAQRPVPTG 241
DB 180 TSDQWILDGPHRGDRLRAGAGANIVPNSGAAKAGIVTPELNGKLDGAAQRPVPTG 239
QY 242 SYTELAVTLDRKNSVDEINAAKKAASNDGFTYEDPIVSSDIVGVSGSLFPAOTQKWE 301
DB 240 SYTELAVTLDRKNSVDEINAAKKAASNDGFTYEDPIVSSDIVGVSGSLFPAOTQKWE 299
QY 302 VDSGQIVKVSNTNENSTYTAQLVETLEPAK 333
DB 300 VDSGQIVKVSNTNENSTYTAQLVETLEPAK 331

RESULT 11
US-09-134-000C-4400
Sequence 4400, Application US/09134000C
Patent No. 6617156

GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 032796-032
CURRENT APPLICATION NUMBER: US/09/134,000C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/055,778
PRIOR FILING DATE: 1997-08-15
NUMBER OF SEQ ID NOS: 6812
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4400
LENGTH: 346
TYPE: PRT
ORGANISM: Enterococcus faecalis
US-09-134-000C-4400

Query Match 78.0%; Score 1337.5; DB 4; Length 346;

Best Local Similarity 78.4%; Pred. No. 4,1e-126;
Matches 262; Conservative 24; Mismatches 47; Indels 1; Gaps 1;

QY 1 MVTXKXINGFGRIQRIAPRIQNVGVYTRINDLTDPMALHLKYDTGGREPTVEV 60
DB 14 MVTXKXINGFGRIQRIAPRIQNVGVYTRINDLTDPMALHLKYDTGGREPTVEV 73
QY 61 KEGGEVNGNFIKVSARDPENIDMATDGEIVLEATGFPFKAKEAABKHLHANGAKKVI 120
DB 74 HEGSFVNGKEIKVLANRNPBELPWGELGVYDLECTGFTFSAAKHLTA-GAKKVI 132
QY 121 TAPGNDVKTVEFNTNHDLDGTETVLSGASCTTNCIAPAKALHDAFGIQGLMTTTHA 180
DB 133 SAPGNDVPTIYVNTNHTLTGKETVLSGASCTTNCIAPAKALHDAFGVSGMTTTHA 192
QY 181 YTGDMILDGPHRGDLRRARAAGANIVPNSGAAKALGLVPELNGKLDGAORVPPT 240
DB 193 YTGDMILDGPHRGDLRRARAAGANIVPNSGAAKALGLVPELNGKLDGAORVPPT 252
QY 241 GSVELEVTLDKGVSVDEINNAKKAASNDGTEPTVSSDIVGSYSLPDAQTCKM 300
DB 253 GSVELEVTLDKGVSVDEINNAKKAASNDGTEPTVSSDIVGSYSLPDAQTCKM 312
QY 301 EVDGSQVTVSVSWYDNEMSYTAQVLTLEFPAKI 334
DB 313 TVGDKQVKTVMVWYDNEMSYTAQVLTLEFPANL 346

RESULT 12

US-09-107-532A-4769
Sequence 4769, Application US/09107532A
Patent No. 6583275

GENERAL INFORMATION:

APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:

ADDRESS: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354

COMPUTER READABLE FORM:

MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Ariniello, Pamela Dereke
REGISTRATION NUMBER: 40,489

TELECOMMUNICATION INFORMATION:
REFERENCE/DOCKET NUMBER: GTC-012
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277

INFORMATION FOR SEQ ID NO: 4769:

SEQUENCE CHARACTERISTICS:
LENGTH: 333 amino acids
TYPE: amino acid

MOLECULE TYPE: protein
HYPOTHEICAL: YES

ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium

FEATURES:
NAME/KEY: misc feature
LOCATION: (8) LOCATION 1...333
SEQUENCE DESCRIPTION: SEQ ID NO: 4769:
US-09-107-532A-4769

Query Match

77.3%; Score 1325.5; DB 4; Length 333;
Best Local Similarity 77.8%; Pred. No. 6.3e-125;
Matches 260; Conservative 25; Mismatches 48; Indels 1; Gaps 1;

QY 1 MVTXKXINGFGRIQRIAPRIQNVGVYTRINDLTDPMALHLKYDTGGREPTVEV 60
DB 1 MVTXKXINGFGRIQRIAPRIQNVGVYTRINDLTDPMALHLKYDTGGREPTVEV 60
QY 61 KEGGEVNGNFIKVSARDPENIDMATDGEIVLEATGFPFKAKEAABKHLHANGAKKVI 120
DB 61 HEGSFVNGKEIKVLANRNPBELPWGELGVYDLECTGFTFSAAKHLTA-GAKKVI 119
QY 121 TAPGNDVKTVEFNTNHDLDGTETVLSGASCTTNCIAPAKALHDAFGIQGLMTTTHA 180
DB 120 SAPGNDVPTIYVNTNHTLTGKETVLSGASCTTNCIAPAKALHDAFGVSGMTTTHA 179
QY 181 YTGDMILDGPHRGDLRRARAAGANIVPNSGAAKALGLVPELNGKLDGAORVPPT 240
DB 180 YTGDMILDGPHRGDLRRARAAGANIVPNSGAAKALGLVPELNGKLDGAORVPPT 229
QY 241 GSVELEVTLDKGVSVDEINNAKKAASNDGTEPTVSSDIVGSYSLPDAQTCKM 300
DB 240 GSVELEVTLDKGVSVDEINNAKKAASNDGTEPTVSSDIVGSYSLPDAQTCKM 299
QY 301 EVDGSQVTVSVSWYDNEMSYTAQVLTLEFPAKI 334
DB 300 TVGDKQVKTVMVWYDNEMSYTAQVLTLEFPANL 333

RESULT 13

US-09-134-001C-5513
Sequence 5513, Application US/09134001C
Patent No. 6380370

GENERAL INFORMATION:

APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCI
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C

CURRENT FILING DATE: 1998-08-13

PRIOR APPLICATION NUMBER: US 60/064,964

PRIOR FILING DATE: 1997-11-08

PRIOR APPLICATION NUMBER: US 60/055,779

PRIOR FILING DATE: 1997-08-14

NUMBER OF SEQ ID NOS: 5674

SEQ ID NO 5513

LENGTH: 348

TYPE: PRT

ORGANISM: Staphylococcus epidermidis

US-09-134-001C-5513

Query Match

69.4%; Score 1190; DB 4; Length 348;
Best Local Similarity 69.4%; Pred. No. 2.9e-111;
Matches 234; Conservative 39; Mismatches 62; Indels 2; Gaps 2;

QY 1 MVTXKXINGFGRIQRIAPRIQNVGVYTRINDLTDPMALHLKYDTGGREPTVEV 60
DB 13 MVTXKXINGFGRIQRIAPRIQNVGVYTRINDLTDPMALHLKYDTGGREPTVEV 72
QY 61 KEGGEVNGNFIKVSARDPENIDMATDGEIVLEATGFPFKAKEAABKHLHANGAKKVI 120
DB 73 HEGSFVNGKEIKVLANRNPBELPWGELGVYDLECTGFTFSAAKHLTA-GAKKVI 131
QY 121 TAPGNDVKTVEFNTNHDLDGTETVLSGASCTTNCIAPAKALHDAFGIQGLMTTTHA 180
DB 132 SAPGNDVPTIYVNTNHTLTGKETVLSGASCTTNCIAPAKALHDAFGVSGMTTTHA 191

QY 181 YTGDMILDPHGGDLRARRAGANIVNSTGAAGAIGLVLPELNGKLDGAORVPY 240
 DB 192 YTGDMILDPHGGDLRARRAGANIVNSTGAAGAIGLVLPELNGKLDGAORVPY 251
 QY 241 GSVTELVTLTK-NVSYDEINAMRAASNDSPGYTEDPIVSSDIVSGSLPDATQYV 299
 DB 252 GSITELTVLTKQVTVDOVNSAMQASDSFGYTEDEIVSSDIQWTSGLPDATQYV 311
 QY 300 MEYDSQVYKVVSVYNDENSYTRAOVRLTEFPAKIAK 336
 DB 312 MTVGDRQVLVXVAAWYNDENSYTRAOVRLTEFPAKIAK 348

RESULT 14

US-09-134-000C-4229
 ; Sequence 4229, Application US/09134C000C
 ; Patent No. 6617156
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn Doucette-Stamm et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
 ; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 032796-032
 ; CURRENT APPLICATION NUMBER: US/09/134,000C
 ; CURRENT FILING DATE: 1998-08-13
 ; PRIOR APPLICATION NUMBER: US 60/055,778
 ; PRIOR FILING DATE: 1997-08-15
 ; NUMBER OF SEQ ID NOS: 6812
 ; SOFTWARE: Patent version 3.1
 ; SEQ ID NO 4229
 ; LENGTH: 357
 ; TYPE: PRT
 ; ORGANISM: Enterococcus faecalis
 ; US-09-134-000C-4229

Query Match 56.4%; Score 967; DB 4; Length 357;
 Best Local Similarity 58.3%; Pred. No. 8e-89;
 Matches 197; Conservative 44; Mismatches 91; Indels 6; Gaps 5;

QY 1 MYKVGINGRIGRLAERLQV--EGEYTRINDLDPNMLAHLYDQSGRFG 59
 DB 22 MYKVGINGRIGRLAERLQV--EGEYTRINDLDPNMLAHLYDQSGRFG 81
 QY 60 VKSGGFVNGNFIKVSARDPENIDWTD-GVEIVLEATGFFAKKEAAEKHLHANGAKY 118
 DB 82 ATENGIVVGEETRYAEPBASKIPWYKENGVDIVLECTGFYSEKKAQALHDA-GVKRY 140
 QY 119 VITAPGNDVKTIVFNTNHDILDGTEYVSGASCTTNCIAPMAKALHDAFGIQGLMTTI 178
 DB 141 VISAPAG-AMKTIYVNVNDTLDNDKIIISAGSCTTNCIAPMAKALHDAFGIQGLMTTI 199
 QY 179 HATGGMILDPHGGDLRARRAGANIVNSTGAAGAIGLVLPELNGKLDGAORVPY 238
 DB 200 HATGGMILDPHGGDLRARRAGANIVNSTGAAGAIGLVLPELNGKLDGAORVPY 259
 QY 229 PTCGVTELVTLTK-NVSYDEINAMRAASNDSPGYTEDPIVSSDIVSGSLPDATQY 296
 DB 260 VDSGTELVTLTK-NVSYDEINAMRAASNDSPGYTEDPIVSSDIVSGSLPDATQY 319
 QY 297 TKYMEVDSQVYKVVSVYNDENSYTRAOVRLTEFPAKIAK 334
 DB 320 TEVTTAGDFQVLVXVAAWYNDENSYTRAOVRLTEFPAKIAK 348

RESULT 15

US-09-634-238-233
 ; Sequence 233, Application US/09634238
 ; Patent No. 6544772
 ; GENERAL INFORMATION:
 ; APPLICANT: Glenn, Matthew
 ; APPLICANT: Hattukala, Ilka J.
 ; APPLICANT: Bloksberg, Leonard, N.
 ; APPLICANT: Lubbers, Mark M.
 ; APPLICANT: Dekker, James

; APPLICANT: Christenson, Anna C.
 ; APPLICANT: Holland, Ross
 ; APPLICANT: O'Toole, Paul W.
 ; APPLICANT: Reid, Julian R.
 ; APPLICANT: Coolbear, Timothy
 ; TITLE OF INVENTION: Polynucleotides, materials incorporating
 ; TITLE OF INVENTION: them and methods for using them.
 ; FILE REFERENCE: 11000.1043U1
 ; CURRENT APPLICATION NUMBER: US/09/634,238
 ; CURRENT FILING DATE: 2000-08-08
 ; NUMBER OF SEQ ID NOS: 422
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 233
 ; LENGTH: 340
 ; TYPE: PRT
 ; ORGANISM: Lactobacillus rhamnosus
 ; US-09-634-238-233

Query Match 56.3%; Score 966; DB 4; Length 340;
 Best Local Similarity 57.0%; Pred. No. 9.4e-89;
 Matches 195; Conservative 48; Mismatches 89; Indels 10; Gaps 6;

QY 1 MYKVGINGRIGRLAERLQV--EGEYTRINDLDPNMLAHLYDQSGRFG 56
 DB 1 MYKVGINGRIGRLAERLQV--EGEYTRINDLDPNMLAHLYDQSGRFG 60
 QY 57 TVEVGGFVNGNFIKVSARDPENIDWTD-GVEIVLEATGFFAKKEAAEKHLHANGAKY 115
 DB 61 EYSATDNQIYVDGKERYAYAEQAQNPWYKNDGVYVIBCGFYTSKESQALHDA-CA 119
 QY 116 KRVVITAPGNDVKTIVFNTNHDILDGTEYVSGASCTTNCIAPMAKALHDAFGIQGLMT 175
 DB 120 KRVVITAPGNDVKTIVFNTNHDILDGTEYVSGASCTTNCIAPMAKALHDAFGIQGLMT 178
 QY 176 TTHAYTQDMILDPHGGDLRARRAGANIVNSTGAAGAIGLVLPELNGKLDGAORVPY 235
 DB 179 TTHAYTQDMILDPHGGDLRARRAGANIVNSTGAAGAIGLVLPELNGKLDGAORVPY 238
 QY 236 VVPTGCVTELVTLTK-NVSYDEINAMRAASNDSPGYTEDPIVSSDIVSGSLPDATQY 292
 DB 239 VVPTGCVTELVTLTK-NVSYDEINAMRAASNDSPGYTEDPIVSSDIVSGSLPDATQY 296
 QY 293 DATGKMEVDSQVYKVVSVYNDENSYTRAOVRLTEFPAKIAK 334
 DB 299 DPTGTEVTTAGDFQVLVXVAAWYNDENSYTRAOVRLTEFPAKIAK 340

Search completed: May 5, 2004, 14:22:18
 Job time: 25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 14:21:14 ; Search time 48 Seconds
(without alignments)
1940.327 Million cell updates/sec

Title: US-09-878-781-4
Perfect score: 1715
Sequence: 1 MVMKVGNGRIGRLAARR.....EMSYRQVLTVEYFAKIAK 336

Scoring table:
BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1138120 seqs, 277189581 residues

Total number of hits satisfying chosen parameters: 1138120

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

Published Applications AA:*
1: /cgm2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgm2_6/ptodata/2/pubpaa/PCT_NEM_PUB.pep:*
3: /cgm2_6/ptodata/2/pubpaa/US06_NEM_PUB.pep:*
4: /cgm2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgm2_6/ptodata/2/pubpaa/US07_NEM_PUB.pep:*
6: /cgm2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgm2_6/ptodata/2/pubpaa/US08_NEM_PUB.pep:*
8: /cgm2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgm2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
10: /cgm2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgm2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
12: /cgm2_6/ptodata/2/pubpaa/US09_PUB.pep:*
13: /cgm2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
14: /cgm2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
15: /cgm2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
16: /cgm2_6/ptodata/2/pubpaa/US10_NEM_PUB.pep:*
17: /cgm2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*
18: /cgm2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1715	100.0	336	9	US-09-878-766A-12
2	1715	100.0	336	10	US-09-878-781-4
3	1715	100.0	336	12	US-10-650-369-12
4	1715	100.0	336	14	US-10-134-297-4
5	1714	99.9	336	12	US-10-282-122A-74379
6	1688.5	98.5	335	10	US-09-878-781-14
7	1688.5	96.6	448	9	US-09-878-766A-22
8	1656.5	96.6	448	12	US-10-650-369-22
9	1618	94.3	336	9	US-09-878-766A-16
10	1618	94.3	336	10	US-09-878-781-8
11	1618	94.3	336	12	US-10-650-369-16
12	1618	94.3	336	14	US-10-134-297-8
13	1611	93.9	336	10	US-09-878-781-16
14	1602	93.4	336	9	US-09-878-766A-14
15	1602	93.4	336	10	US-09-878-781-6

16	1602	93.4	336	12	US-10-650-369-14	Sequence 14, Appl
17	1602	93.4	336	14	US-10-134-297-6	Sequence 6, Appl
18	1593.5	92.9	359	9	US-09-815-242-13169	Sequence 13169, A
19	1593.5	92.9	359	12	US-10-282-122A-74186	Sequence 74186, A
20	1592.5	92.9	359	9	US-09-815-242-13593	Sequence 13593, A
21	1584	92.4	336	9	US-09-878-766A-20	Sequence 20, Appl
22	1584	92.4	336	10	US-09-878-781-12	Sequence 12, Appl
23	1584	92.4	336	12	US-10-650-369-20	Sequence 20, Appl
24	1584	92.4	336	14	US-10-134-297-12	Sequence 12, Appl
25	1579.5	92.1	337	12	US-10-282-122A-72254	Sequence 72254, A
26	1578	92.0	336	9	US-09-878-766A-18	Sequence 18, Appl
27	1578	92.0	336	10	US-09-878-781-10	Sequence 10, Appl
28	1578	92.0	336	12	US-10-650-369-18	Sequence 18, Appl
29	1578	92.0	336	14	US-10-134-297-10	Sequence 10, Appl
30	1549.5	90.3	333	9	US-09-765-272-54	Sequence 54, Appl
31	1410	82.2	336	15	US-10-369-493-18524	Sequence 18524, A
32	1338.5	78.0	337	15	US-10-369-493-18342	Sequence 18342, A
33	1337.5	78.0	333	9	US-09-815-242-10847	Sequence 10847, A
34	1337.5	78.0	333	12	US-10-282-122A-57294	Sequence 57294, A
35	1277	74.5	335	12	US-10-282-122A-53266	Sequence 53266, A
36	1262	73.6	336	12	US-10-282-122A-51519	Sequence 51519, A
37	1259.5	73.4	334	12	US-10-282-122A-71734	Sequence 71734, A
38	1217.5	71.0	335	12	US-10-282-122A-65661	Sequence 65661, A
39	1197.5	69.8	334	12	US-10-282-122A-65001	Sequence 65001, A
40	1196.5	69.8	357	12	US-10-282-122A-70890	Sequence 70890, A
41	1190	69.4	336	12	US-10-282-122A-5618	Sequence 5618, Ap
42	1183	69.0	336	9	US-09-815-242-12540	Sequence 12540, A
43	1183	69.0	336	12	US-10-282-122A-43997	Sequence 43997, A
44	1183	69.0	336	12	US-10-282-122A-56828	Sequence 56828, A
45	993	57.9	333	12	US-10-282-122A-56828	Sequence 56828, A

ALIGNMENTS

RESULT 1
US-09-878-766A-12 Application US/09878766A
; Sequence 12, Application US/09878766A
; Patent No. US20020044928A1
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION
; PILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/09/878, 766A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus dysgalactiae
US-09-878-766A-12

Query Match 100.0%; Score 1715; DB 9; Length 336;
Best Local Similarity 100.0%; Pred. No. 3.1e-154;

Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVMKVGNGRIGRLAARR.....EMSYRQVLTVEYFAKIAK 60
DB 1 MVMKVGNGRIGRLAARR.....EMSYRQVLTVEYFAKIAK 60
QY 61 KEGGEVNGFIVSAERPDENIDAGVLEATGFFAKKAERKHLNAAKAKVVI 120
DB 61 KEGGEVNGFIVSAERPDENIDAGVLEATGFFAKKAERKHLNAAKAKVVI 120
QY 121 TAPGNDVKVVTNTNHDITGTEIVTSGASTTNCAPMAKALHDFGQKGMATTTHA 180
DB 121 TAPGNDVKVVTNTNHDITGTEIVTSGASTTNCAPMAKALHDFGQKGMATTTHA 180
QY 181 YTGQKHLIDGPHGSGDLRABAGAAIVPNSITGAAKAGLVLPETNGKLDGAAQRPVPT 240

Db 181 YTGOMLIDGPHRGDILRRAGAANIVPNSGAKAIGLIVPELNGKLDGAARVPPT 240
Qy 241 GSVELVVTLDKNVSVDEINAAKKAASNDSPGYTDEPIVSSDIVGVSYSGLFDATQTKVM 300
Db 241 GSVELVVTLDKNVSVDEINAAKKAASNDSPGYTDEPIVSSDIVGVSYSGLFDATQTKVM 300
Qy 301 EVDGSQLVXVSWYDNEMSYTAOLVETLEFPAKIAK 336
Db 301 EVDGSQLVXVSWYDNEMSYTAOLVETLEFPAKIAK 336

RESULT 2

US-09-878-781-4
Sequence 4, Application US/09878781
Publication No. US20030082781A1
GENERAL INFORMATION:
APPLICANT: Bolton, Alexandra J.
APPLICANT: Perez-Casal, Jose
APPLICANT: Fontaine, Michael
APPLICANT: Potter, Andrew A.
TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GAPC PROTEIN AGAINST
FILE REFERENCE: 9000-0055
CURRENT APPLICATION NUMBER: US/09/878,781
CURRENT FILING DATE: 2002-09-10
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 336
TYPE: PRT
ORGANISM: Streptococcus dysgalactiae
US-09-878-781-4

Query Match 100.0%; Score 1715; DB 10; Length 336;
Best Local Similarity 100.0%; Pred. No. 3,1e-154;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MYVKGINGSGRIGRLAFRIQNVGVEVTRINDLDPNMLAHLLKYDTTQGRFDGVEV 60
Db 1 MYVKGINGSGRIGRLAFRIQNVGVEVTRINDLDPNMLAHLLKYDTTQGRFDGVEV 60
Qy 61 KEGGEVNGNFIVKSAERDPENIDWATDGEIVLEATGFFAKKEAAEKHLHANGAKKVV 120
Db 61 KEGGEVNGNFIVKSAERDPENIDWATDGEIVLEATGFFAKKEAAEKHLHANGAKKVV 120
Qy 121 TAPGANDVKTVPVNTNHDILDGTEVVISGASCTTNCLAPMAKALHDARFGIQKGLMTTHA 180
Db 121 TAPGANDVKTVPVNTNHDILDGTEVVISGASCTTNCLAPMAKALHDARFGIQKGLMTTHA 180
Qy 181 YTGDOMLIDGPHRGDILRRAGAANIVPNSGAKAIGLIVPELNGKLDGAARVPPT 240
Db 181 YTGDOMLIDGPHRGDILRRAGAANIVPNSGAKAIGLIVPELNGKLDGAARVPPT 240
Qy 241 GSVELVVTLDKNVSVDEINAAKKAASNDSPGYTDEPIVSSDIVGVSYSGLFDATQTKVM 300
Db 241 GSVELVVTLDKNVSVDEINAAKKAASNDSPGYTDEPIVSSDIVGVSYSGLFDATQTKVM 300
Qy 301 EVDGSQLVXVSWYDNEMSYTAOLVETLEFPAKIAK 336
Db 301 EVDGSQLVXVSWYDNEMSYTAOLVETLEFPAKIAK 336

RESULT 3

US-10-650-369-12
Sequence 12, Application US/10650369
Publication No. US20040062774A1
GENERAL INFORMATION:
APPLICANT: Potter, Andrew A.
APPLICANT: Perez-Casal, Jose
APPLICANT: Fontaine, Michael
TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
AGAINST STREPTOCOCCUS INFECTION

FILE REFERENCE: 9000-0057
CURRENT APPLICATION NUMBER: US/10/650,369
CURRENT FILING DATE: 2003-08-27
PRIOR APPLICATION NUMBER: US/09/878,786A
PRIORITY FILING DATE: 2001-09-10
NUMBER OF SEQ ID NOS: 22
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 12
LENGTH: 336
TYPE: PRT
ORGANISM: Streptococcus dysgalactiae
US-10-650-369-12

Query Match 100.0%; Score 1715; DB 12; Length 336;
Best Local Similarity 100.0%; Pred. No. 3,1e-154;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MYVKGINGSGRIGRLAFRIQNVGVEVTRINDLDPNMLAHLLKYDTTQGRFDGVEV 60
Db 1 MYVKGINGSGRIGRLAFRIQNVGVEVTRINDLDPNMLAHLLKYDTTQGRFDGVEV 60
Qy 61 KEGGEVNGNFIVKSAERDPENIDWATDGEIVLEATGFFAKKEAAEKHLHANGAKKVV 120
Db 61 KEGGEVNGNFIVKSAERDPENIDWATDGEIVLEATGFFAKKEAAEKHLHANGAKKVV 120
Qy 121 TAPGANDVKTVPVNTNHDILDGTEVVISGASCTTNCLAPMAKALHDARFGIQKGLMTTHA 180
Db 121 TAPGANDVKTVPVNTNHDILDGTEVVISGASCTTNCLAPMAKALHDARFGIQKGLMTTHA 180
Qy 181 YTGDOMLIDGPHRGDILRRAGAANIVPNSGAKAIGLIVPELNGKLDGAARVPPT 240
Db 181 YTGDOMLIDGPHRGDILRRAGAANIVPNSGAKAIGLIVPELNGKLDGAARVPPT 240
Qy 241 GSVELVVTLDKNVSVDEINAAKKAASNDSPGYTDEPIVSSDIVGVSYSGLFDATQTKVM 300
Db 241 GSVELVVTLDKNVSVDEINAAKKAASNDSPGYTDEPIVSSDIVGVSYSGLFDATQTKVM 300
Qy 301 EVDGSQLVXVSWYDNEMSYTAOLVETLEFPAKIAK 336
Db 301 EVDGSQLVXVSWYDNEMSYTAOLVETLEFPAKIAK 336

RESULT 4

US-10-134-297-4
Sequence 4, Application US/10134297
Publication No. US2003016524A1
GENERAL INFORMATION:
APPLICANT: Bolton, Alexandra J.
APPLICANT: Perez-Casal, Jose
APPLICANT: Fontaine, Michael
APPLICANT: Potter, Andrew A.
TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GAPC PROTEIN AGAINST
FILE REFERENCE: 9000-0055.20
CURRENT APPLICATION NUMBER: US/10/134,297
CURRENT FILING DATE: 2002-04-26
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 336
TYPE: PRT
ORGANISM: Streptococcus dysgalactiae
US-10-134-297-4

Query Match 100.0%; Score 1715; DB 14; Length 336;
Best Local Similarity 100.0%; Pred. No. 3,1e-154;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MYVKGINGSGRIGRLAFRIQNVGVEVTRINDLDPNMLAHLLKYDTTQGRFDGVEV 60
Db 1 MYVKGINGSGRIGRLAFRIQNVGVEVTRINDLDPNMLAHLLKYDTTQGRFDGVEV 60
Qy 61 KEGGEVNGNFIVKSAERDPENIDWATDGEIVLEATGFFAKKEAAEKHLHANGAKKVV 120

```
Db 61 KEGGFEVNGNFIVKVAERDPENIDMAITDGEIVLEATGFAKKEAERKHLHANGAKKVI 120
Qy 121 TAPGNDVKTIVFNTNHDILDTETVLSGASCTTNCIAPAAKALHDAFGIQKGLMTTIIA 180
Db 121 TAPGNDVKTIVFNTNHDILDTETVLSGASCTTNCIAPAAKALHDAFGIQKGLMTTIIA 180
Qy 181 YTGDMILDGPHRGDILRRARAGANIVPNSGAAKALIGLVIPELNGKLDGAQRVPPT 240
Db 181 YTGDMILDGPHRGDILRRARAGANIVPNSGAAKALIGLVIPELNGKLDGAQRVPPT 240
Qy 241 GSVTELVTLIDKQVSVDEINAAKASNDSPYEDPIVSSDIIVGVSGLFPAQTQTKM 300
Db 241 GSVTELVTLIDKQVSVDEINAAKASNDSPYEDPIVSSDIIVGVSGLFPAQTQTKM 300
Qy 301 EVDGSQLVKKVSWYDNEMSYTAQIVTLEYFAKIAK 336
Db 301 EVDGSQLVKKVSWYDNEMSYTAQIVTLEYFAKIAK 336
```

RESULT 5

```
US-10-282-122A-74379
; Sequence 74379, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Foreyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: EUTRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; Remaining Prior Application data removed - See File Wrapper or PAM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 74379
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes
US-10-282-122A-74379
```

```
Query Match 99.9%; Score 1714; DB 12; Length 336;
Best Local Similarity 99.7%; Pred. No. 3.8e-154;
Matches 335; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 MYVVGINGRIGRLAFLRRIQNEVEVTRINDITDPNMLAHLLKYDTTQGRFGCTVEV 60
Db 1 MYVVGINGRIGRLAFLRRIQNEVEVTRINDITDPNMLAHLLKYDTTQGRFGCTVEV 60
Qy 61 KEGGFEVNGNFIVKVAERDPENIDMAITDGEIVLEATGFAKKEAERKHLHANGAKKVI 120
Db 61 KEGGFEVNGNFIVKVAERDPENIDMAITDGEIVLEATGFAKKEAERKHLHANGAKKVI 120
Qy 121 TAPGNDVKTIVFNTNHDILDTETVLSGASCTTNCIAPAAKALHDAFGIQKGLMTTIIA 180
Db 121 TAPGNDVKTIVFNTNHDILDTETVLSGASCTTNCIAPAAKALHDAFGIQKGLMTTIIA 180
Qy 181 YTGDMILDGPHRGDILRRARAGANIVPNSGAAKALIGLVIPELNGKLDGAQRVPPT 240
Db 181 YTGDMILDGPHRGDILRRARAGANIVPNSGAAKALIGLVIPELNGKLDGAQRVPPT 240
Qy 241 GSVTELVTLIDKQVSVDEINAAKASNDSPYEDPIVSSDIIVGVSGLFPAQTQTKM 300
Db 241 GSVTELVTLIDKQVSVDEINAAKASNDSPYEDPIVSSDIIVGVSGLFPAQTQTKM 300
Qy 301 EVDGSQLVKKVSWYDNEMSYTAQIVTLEYFAKIAK 336
Db 301 EVDGSQLVKKVSWYDNEMSYTAQIVTLEYFAKIAK 336
```

RESULT 6

```
US-09-878-781-14
; Sequence 14, Application US/09878781
; Publication No. US20030082781A1
; GENERAL INFORMATION:
; APPLICANT: Bolton, Alexandra J.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; APPLICANT: Potter, Andrew A.
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GAPC PROTEIN AGAINST
; FILE REFERENCE: 9000-0055
; CURRENT APPLICATION NUMBER: US/09/878,781
; CURRENT FILING DATE: 2002-09-10
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SpyGAPC
; OTHER INFORMATION: protein
US-09-878-781-14
```

```
Query Match 98.5%; Score 1688.5; DB 10; Length 335;
Best Local Similarity 98.8%; Pred. No. 1e-151;
Matches 332; Conservative 2; Mismatches 1; Indels 1; Gaps 1;
```

```
Qy 1 MYVVGINGRIGRLAFLRRIQNEVEVTRINDITDPNMLAHLLKYDTTQGRFGCTVEV 60
Db 1 MYVVGINGRIGRLAFLRRIQNEVEVTRINDITDPNMLAHLLKYDTTQGRFGCTVEV 60
Qy 61 KEGGFEVNGNFIVKVAERDPENIDMAITDGEIVLEATGFAKKEAERKHLHANGAKKVI 120
Db 61 KEGGFEVNGNFIVKVAERDPENIDMAITDGEIVLEATGFAKKEAERKHLHANGAKKVI 119
Qy 121 TAPGNDVKTIVFNTNHDILDTETVLSGASCTTNCIAPAAKALHDAFGIQKGLMTTIIA 180
Db 121 TAPGNDVKTIVFNTNHDILDTETVLSGASCTTNCIAPAAKALHDAFGIQKGLMTTIIA 179
Qy 181 YTGDMILDGPHRGDILRRARAGANIVPNSGAAKALIGLVIPELNGKLDGAQRVPPT 240
Db 181 YTGDMILDGPHRGDILRRARAGANIVPNSGAAKALIGLVIPELNGKLDGAQRVPPT 239
Qy 241 GSVTELVTLIDKQVSVDEINAAKASNDSPYEDPIVSSDIIVGVSGLFPAQTQTKM 300
Db 241 GSVTELVTLIDKQVSVDEINAAKASNDSPYEDPIVSSDIIVGVSGLFPAQTQTKM 300
```


Dh 240 GSVTELVTLIDKNVSVDEINSAKMAKASNDSPGYEDPIVSSDIVGSYSLDPATQTKM 299
Qy 301 EVDGSQLVYKVVSWYDNEMSYTAQVLTLEYFAKIAK 336
Dh 300 EVDGSQLVYKVVSWYDNEMSYTAQVLTLEYFAKIAK 335

RESULT 7

US-09-878-766A-22
; Sequence 22, Application US/09878766A
; Patent No. US20020044928A1
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/09/878, 766A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 448
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Gapc multiple
; OTHER INFORMATION: epitope fusion protein
US-09-878-766A-22

Query Match 96.6%; Score 1656.5; DB 9; Length 448;
Best Local Similarity 79.6%; Pred. No. 1,7e-148;
Matches 335; Conservative 0; Mismatches 1; Indels 85; Gaps 1;

Qy 1 MYYKVGINGFRIGRLAFRIQNVGEVTRINDLTDPMALHLKYDTTQGRPDGTYEV 60
Dh 28 MYYKVGINGFRIGRLAFRIQNVGEVTRINDLTDPMALHLKYDTTQGRPDGTYEV 87
Qy 61 KEGGFVNKNFIKYSARDE----- 81
Dh 88 KEGGFVNKNFIKYSARDEPNIDMTDGVIEVLEALGTEVKGSGFDVNGKFIKYSAE 147
Qy 62 -----NIDMTDGVIEVLE 95
Dh 148 KDEQIDMTDGVIEVLEIDGTEVKEGGEFVNQPFVKSAREBPANIDMTDGVIEVLE 207
Qy 96 ATGFPAKKAERKHLHANGAKKVVITAPGNDVTVVFNTHDILDGETVYISASCTTN 155
Dh 208 ATGFPAKKAERKHLHANGAKKVVITAPGNDVTVVFNTHDILDGETVYISASCTTN 267
Qy 156 CLAPMAKALHDAFGIQGLMTTHAYTGDMILDGPHRGDLRRARAGANIVPNSTGA 215
Dh 268 CLAPMAKALHDAFGIQGLMTTHAYTGDMILDGPHRGDLRRARAGANIVPNSTGA 327
Qy 216 KAIGVIPELNGKLDGAQRVVPPTGSTEIVTLIDKNVSVDEINAAKMAASNDSPGYTE 275
Dh 328 KAIGVIPELNGKLDGAQRVVPPTGSTEIVTLIDKNVSVDEINAAKMAASNDSPGYTE 387
Qy 276 DPIVSSDIVGSYSLDPATQTKMEVDGSQLVYKVVSWYDNEMSYTAQVLTLEYFAKIA 335
Dh 388 DPIVSSDIVGSYSLDPATQTKMEVDGSQLVYKVVSWYDNEMSYTAQVLTLEYFAKIA 447
Qy 336 K 336
Dh 448 K 448

RESULT 8

US-10-650-369-22
; Sequence 22, Application US/10650369
; Publication No. US2004006277A1
; GENERAL INFORMATION:

; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/10/650, 369
; CURRENT FILING DATE: 2003-08-27
; PRIOR APPLICATION NUMBER: US/09/878, 766A
; PRIOR FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 448
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Gapc multiple
; OTHER INFORMATION: epitope fusion protein
US-10-650-369-22

Query Match 96.6%; Score 1656.5; DB 12; Length 448;
Best Local Similarity 79.6%; Pred. No. 1,7e-148;
Matches 335; Conservative 0; Mismatches 1; Indels 85; Gaps 1;

Qy 1 MYYKVGINGFRIGRLAFRIQNVGEVTRINDLTDPMALHLKYDTTQGRPDGTYEV 60
Dh 28 MYYKVGINGFRIGRLAFRIQNVGEVTRINDLTDPMALHLKYDTTQGRPDGTYEV 87
Qy 61 KEGGFVNKNFIKYSARDE----- 81
Dh 88 KEGGFVNKNFIKYSARDEPNIDMTDGVIEVLEALGTEVKGSGFDVNGKFIKYSAE 147
Qy 82 -----NIDMTDGVIEVLE 95
Dh 148 KDEQIDMTDGVIEVLEIDGTEVKEGGEFVNQPFVKSAREBPANIDMTDGVIEVLE 207
Qy 96 ATGFPAKKAERKHLHANGAKKVVITAPGNDVTVVFNTHDILDGETVYISASCTTN 155
Dh 208 ATGFPAKKAERKHLHANGAKKVVITAPGNDVTVVFNTHDILDGETVYISASCTTN 267
Qy 156 CLAPMAKALHDAFGIQGLMTTHAYTGDMILDGPHRGDLRRARAGANIVPNSTGA 215
Dh 268 CLAPMAKALHDAFGIQGLMTTHAYTGDMILDGPHRGDLRRARAGANIVPNSTGA 327
Qy 216 KAIGVIPELNGKLDGAQRVVPPTGSTEIVTLIDKNVSVDEINAAKMAASNDSPGYTE 275
Dh 328 KAIGVIPELNGKLDGAQRVVPPTGSTEIVTLIDKNVSVDEINAAKMAASNDSPGYTE 387
Qy 276 DPIVSSDIVGSYSLDPATQTKMEVDGSQLVYKVVSWYDNEMSYTAQVLTLEYFAKIA 335
Dh 388 DPIVSSDIVGSYSLDPATQTKMEVDGSQLVYKVVSWYDNEMSYTAQVLTLEYFAKIA 447
Qy 336 K 336
Dh 448 K 448

RESULT 9

US-09-878-766A-16
; Sequence 16, Application US/09878766A
; Patent No. US20020044928A1
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/09/878, 766A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 16
LENGTH: 336
TYPE: PR
ORGANISM: Streptococcus uberis
US-09-878-766A-16

Query Match 94.3%; Score 1618; DB 9; Length 336;
Best Local Similarity 92.9%; Pred. No. 5e-145;
Matches 312; Conservative 15; Mismatches 9; Indels 0; Gaps 0;

QY 1 M V V K G I N G F R I G R L A F R I O N V E G V E T R I N D L T D P M L A H L K Y D T T O G R F D G T V E V 60
DB 1 M V V K G I N G F R I G R L A F R I O N V E G V E T R I N D L T D P M L A H L K Y D T T O G R F D G T V E V 60
QY 61 K G G F E V N G N F I K V S A R D P E N I D M A T D G V E I V L E A T G F A K K A A E K H L H A N G A K V Y I 120
DB 61 K G G F E V N G N F I K V S A R D P E N I D M A T D G V E I V L E A T G F A K K A A E K H L H A N G A K V Y I 120
QY 121 T A P G N D V K T V F E N T N H D I L D G T E T V I S G A S C T T N C L A P M A K L Q D N F G V K G S L P D A T O T K M 180
DB 121 T A P G N D V K T V F E N T N H D I L D G T E T V I S G A S C T T N C L A P M A K L Q D N F G V K G S L P D A T O T K M 180
QY 181 Y T G D M I L D G P H R G D L R R A R A G A N I V N S T G A A A I G L V I P E L N G K L D G A A G R V P P T 240
DB 181 Y T G D M I L D G P H R G D L R R A R A G A N I V N S T G A A A I G L V I P E L N G K L D G A A G R V P P T 240
QY 241 G S Y T E L V A V L E K E T S V E R I N A M K A A N D S Y G T E D P I V S S D I I G M A Y S L P D A T O T K M 300
DB 241 G S Y T E L V A V L E K E T S V E R I N A M K A A N D S Y G T E D P I V S S D I I G M A Y S L P D A T O T K M 300
QY 301 E V D S G Q L V K V S W Y D N E M S Y T A Q L V R T L E Y F A K I A K 336
DB 301 T V D G N Q L V K V S W Y D N E M S Y T A Q L V R T L E Y F A K I A K 336

RESULT 10

US-09-878-781-8
Sequence 8, Application US/09878781
Publication No. US20030082781A1
GENERAL INFORMATION:

APPLICANT: Bolton, Alexandra J.
APPLICANT: Perez-Casal, Jose
APPLICANT: Fontaine, Michael
APPLICANT: Potter, Andrew A.
TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GAPC PROTEIN AGAINST
TITLE OF INVENTION: STREPTOCOCCUS INFECTION
FILE REFERENCE: 9000-0055
CURRENT APPLICATION NUMBER: US/09/878, 781
CURRENT FILING DATE: 2002-09-10
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 8
LENGTH: 336
TYPE: PR
ORGANISM: Streptococcus uberis
US-09-878-781-8

Query Match 94.3%; Score 1619; DB 10; Length 336;
Best Local Similarity 92.9%; Pred. No. 5e-145;
Matches 312; Conservative 15; Mismatches 9; Indels 0; Gaps 0;

QY 1 M V V K G I N G F R I G R L A F R I O N V E G V E T R I N D L T D P M L A H L K Y D T T O G R F D G T V E V 60
DB 1 M V V K G I N G F R I G R L A F R I O N V E G V E T R I N D L T D P M L A H L K Y D T T O G R F D G T V E V 60
QY 61 K G G F E V N G N F I K V S A R D P E N I D M A T D G V E I V L E A T G F A K K A A E K H L H A N G A K V Y I 120
DB 61 K G G F E V N G N F I K V S A R D P E N I D M A T D G V E I V L E A T G F A K K A A E K H L H A N G A K V Y I 120
QY 121 T A P G N D V K T V F E N T N H D I L D G T E T V I S G A S C T T N C L A P M A K L Q D N F G V K G S L P D A T O T K M 180
DB 121 T A P G N D V K T V F E N T N H D I L D G T E T V I S G A S C T T N C L A P M A K L Q D N F G V K G S L P D A T O T K M 180

QY 181 Y T G D M I L D G P H R G D L R R A R A G A N I V N S T G A A A I G L V I P E L N G K L D G A A G R V P P T 240
DB 181 Y T G D M I L D G P H R G D L R R A R A G A N I V N S T G A A A I G L V I P E L N G K L D G A A G R V P P T 240
QY 241 G S Y T E L V A V L E K E T S V E R I N A M K A A N D S Y G T E D P I V S S D I I G M A Y S L P D A T O T K M 300
DB 241 G S Y T E L V A V L E K E T S V E R I N A M K A A N D S Y G T E D P I V S S D I I G M A Y S L P D A T O T K M 300
QY 301 E V D S G Q L V K V S W Y D N E M S Y T A Q L V R T L E Y F A K I A K 336
DB 301 T V D G N Q L V K V S W Y D N E M S Y T A Q L V R T L E Y F A K I A K 336

RESULT 11

US-10-650-369-16
Sequence 16, Application US/10650369
Publication No. US20040062774A1
GENERAL INFORMATION:

APPLICANT: Potter, Andrew A.
APPLICANT: Perez-Casal, Jose
APPLICANT: Fontaine, Michael
TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION
FILE REFERENCE: 9000-0057
CURRENT APPLICATION NUMBER: US/10/650,369
CURRENT FILING DATE: 2003-08-27
PRIOR APPLICATION NUMBER: US/09/878, 766A
PRIOR FILING DATE: 2001-09-10
NUMBER OF SEQ ID NOS: 22
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 16
LENGTH: 336
TYPE: PR
ORGANISM: Streptococcus uberis
US-10-650-369-16

Query Match 94.3%; Score 1618; DB 12; Length 336;
Best Local Similarity 92.9%; Pred. No. 5e-145;
Matches 312; Conservative 15; Mismatches 9; Indels 0; Gaps 0;

QY 1 M V V K G I N G F R I G R L A F R I O N V E G V E T R I N D L T D P M L A H L K Y D T T O G R F D G T V E V 60
DB 1 M V V K G I N G F R I G R L A F R I O N V E G V E T R I N D L T D P M L A H L K Y D T T O G R F D G T V E V 60
QY 61 K G G F E V N G N F I K V S A R D P E N I D M A T D G V E I V L E A T G F A K K A A E K H L H A N G A K V Y I 120
DB 61 K G G F E V N G N F I K V S A R D P E N I D M A T D G V E I V L E A T G F A K K A A E K H L H A N G A K V Y I 120
QY 121 T A P G N D V K T V F E N T N H D I L D G T E T V I S G A S C T T N C L A P M A K L Q D N F G V K G S L P D A T O T K M 180
DB 121 T A P G N D V K T V F E N T N H D I L D G T E T V I S G A S C T T N C L A P M A K L Q D N F G V K G S L P D A T O T K M 180
QY 181 Y T G D M I L D G P H R G D L R R A R A G A N I V N S T G A A A I G L V I P E L N G K L D G A A G R V P P T 240
DB 181 Y T G D M I L D G P H R G D L R R A R A G A N I V N S T G A A A I G L V I P E L N G K L D G A A G R V P P T 240
QY 241 G S Y T E L V A V L E K E T S V E R I N A M K A A N D S Y G T E D P I V S S D I I G M A Y S L P D A T O T K M 300
DB 241 G S Y T E L V A V L E K E T S V E R I N A M K A A N D S Y G T E D P I V S S D I I G M A Y S L P D A T O T K M 300
QY 301 E V D S G Q L V K V S W Y D N E M S Y T A Q L V R T L E Y F A K I A K 336
DB 301 T V D G N Q L V K V S W Y D N E M S Y T A Q L V R T L E Y F A K I A K 336

RESULT 12

US-10-134-297-8
Sequence 8, Application US/10134297
Publication No. US20030165524A1
GENERAL INFORMATION:

APPLICANT: Bolton, Alexandra J.
APPLICANT: Perez-Casal, Jose
APPLICANT: Fontaine, Michael

APPLICANT: Potter, Andrew A.
TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GAPC PROTEIN AGAINST
FILE REFERENCE: 9600-0055.20
CURRENT APPLICATION NUMBER: US/10/134.297
CURRENT FILING DATE: 2002-04-26
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 8
LENGTH: 336
TYPE: PRT
ORGANISM: Streptococcus uberis
US-10-134-297-8

Query Match: 94.3%; Score 1619; DB 14; Length 336;
Best Local Similarity 92.8%; Pred. No. 5e-145;
Matches 312; Conservative 15; Mismatches 9; Indels 0; Gaps 0;

QY 1 MVRKVGINGFGRIGRLAFRIQNVGEVETRIINDLTPNMLAHLLKXTDTTQGRFDGTEV 60
DB 1 MVRKVGINGFGRIGRLAFRIQNVGEVETRIINDLTPNMLAHLLKXTDTTQGRFDGTEV 60
QY 61 KEGFEVKNPFIKVAERDPENIDMTDGEIVLEATGFPKKAABKHLHANGAKKVI 120
DB 61 KEGFEVKNPFIKVAERDPENIDMTDGEIVLEATGFPKKAABKHLHANGAKKVI 120
QY 121 TAPGNDVKTIVFNNHDIIDGTEIVISGASCTTNCIAPAKKALHDAFCIQKGLMTTIIA 180
DB 121 TAPGNDVKTIVFNNHDIIDGTEIVISGASCTTNCIAPAKKALHDAFCIQKGLMTTIIA 180
QY 181 YTGDMITDQPHRGDILRRARAGANIVPNSGAKAIGVITPELNGKLDGAAGRPVPT 240
DB 181 YTGDMITDQPHRGDILRRARAGANIVPNSGAKAIGVITPELNGKLDGAAGRPVPT 240
QY 241 GSVTELVVTLIDKRVSVDEINAMKAASNDSPGTEDPIVSSDIIVGVSYSGLFDATQTKM 300
DB 241 GSVTELVVTLIDKRVSVDEINAMKAASNDSPGTEDPIVSSDIIVGVSYSGLFDATQTKM 300
QY 301 EVDGSQLVKVVSMYDNEMSYTAQVLTLEYFAKIAK 336
DB 301 EVDGSQLVKVVSMYDNEMSYTAQVLTLEYFAKIAK 336

RESULT 13
US-09-878-781-16
Sequence 16, Application US/09878781
Publication No. US20030082781A1
GENERAL INFORMATION:
APPLICANT: Bolton, Alexandra J.
APPLICANT: Perez-Casal, Jose
APPLICANT: Fontaine, Michael
APPLICANT: Potter, Andrew A.
TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GAPC PROTEIN AGAINST
FILE REFERENCE: 9000-0055
CURRENT APPLICATION NUMBER: US/09/878,781
CURRENT FILING DATE: 2002-09-10
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 16
LENGTH: 336
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURES:
OTHER INFORMATION: Description of Artificial Sequence: SeqGapC
US-09-878-781-16

Query Match: 93.9%; Score 1611; DB 10; Length 336;
Best Local Similarity 95.2%; Pred. No. 2.3e-144;
Matches 320; Conservative 2; Mismatches 14; Indels 0; Gaps 0;
QY 1 MVRKVGINGFGRIGRLAFRIQNVGEVETRIINDLTPNMLAHLLKXTDTTQGRFDGTEV 60

DB 1 MVRKVGINGFGRIGRLAFRIQNVGEVETRIINDLTPNMLAHLLKXTDTTQGRFDGTEV 60
QY 61 KEGFEVKNPFIKVAERDPENIDMTDGEIVLEATGFPKKAABKHLHANGAKKVI 120
DB 61 KEGFEVKNPFIKVAERDPENIDMTDGEIVLEATGFPKKAABKHLHANGAKKVI 120
QY 121 TAPGNDVKTIVFNNHDIIDGTEIVISGASCTTNCIAPAKKALHDAFCIQKGLMTTIIA 180
DB 121 TAPGNDVKTIVFNNHDIIDGTEIVISGASCTTNCIAPAKKALHDAFCIQKGLMTTIIA 180
QY 181 YTGDMITDQPHRGDILRRARAGANIVPNSGAKAIGVITPELNGKLDGAAGRPVPT 240
DB 181 YTGDMITDQPHRGDILRRARAGANIVPNSGAKAIGVITPELNGKLDGAAGRPVPT 240
QY 241 GSVTELVVTLIDKRVSVDEINAMKAASNDSPGTEDPIVSSDIIVGVSYSGLFDATQTKM 300
DB 241 GSVTELVVTLIDKRVSVDEINAMKAASNDSPGTEDPIVSSDIIVGVSYSGLFDATQTKM 300
QY 301 EVDGSQLVKVVSMYDNEMSYTAQVLTLEYFAKIAK 336
DB 301 EVDGSQLVKVVSMYDNEMSYTAQVLTLEYFAKIAK 336

RESULT 14
US-09-878-766A-14
Sequence 14, Application US/09878766A
Patent No. US20020044928A1
GENERAL INFORMATION:
APPLICANT: Potter, Andrew A.
APPLICANT: Perez-Casal, Jose
APPLICANT: Fontaine, Michael
TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
FILE REFERENCE: 9000-0057
CURRENT APPLICATION NUMBER: US/09/878,766A
CURRENT FILING DATE: 2001-09-10
NUMBER OF SEQ ID NOS: 22
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 14
LENGTH: 336
TYPE: PRT
ORGANISM: Streptococcus agalactiae
US-09-878-766A-14

Query Match: 93.4%; Score 1602; DB 9; Length 336;
Best Local Similarity 91.7%; Pred. No. 1.7e-143;
Matches 308; Conservative 17; Mismatches 11; Indels 0; Gaps 0;

QY 1 MVRKVGINGFGRIGRLAFRIQNVGEVETRIINDLTPNMLAHLLKXTDTTQGRFDGTEV 60
DB 1 MVRKVGINGFGRIGRLAFRIQNVGEVETRIINDLTPNMLAHLLKXTDTTQGRFDGTEV 60
QY 61 KEGFEVKNPFIKVAERDPENIDMTDGEIVLEATGFPKKAABKHLHANGAKKVI 120
DB 61 KEGFEVKNPFIKVAERDPENIDMTDGEIVLEATGFPKKAABKHLHANGAKKVI 120
QY 121 TAPGNDVKTIVFNNHDIIDGTEIVISGASCTTNCIAPAKKALHDAFCIQKGLMTTIIA 180
DB 121 TAPGNDVKTIVFNNHDIIDGTEIVISGASCTTNCIAPAKKALHDAFCIQKGLMTTIIA 180
QY 181 YTGDMITDQPHRGDILRRARAGANIVPNSGAKAIGVITPELNGKLDGAAGRPVPT 240
DB 181 YTGDMITDQPHRGDILRRARAGANIVPNSGAKAIGVITPELNGKLDGAAGRPVPT 240
QY 241 GSVTELVVTLIDKRVSVDEINAMKAASNDSPGTEDPIVSSDIIVGVSYSGLFDATQTKM 300
DB 241 GSVTELVVTLIDKRVSVDEINAMKAASNDSPGTEDPIVSSDIIVGVSYSGLFDATQTKM 300
QY 301 EVDGSQLVKVVSMYDNEMSYTAQVLTLEYFAKIAK 336
DB 301 EVDGSQLVKVVSMYDNEMSYTAQVLTLEYFAKIAK 336

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 14:17:44 ; Search time 20 Seconds

(without alignments)
1616.018 Million cell updates/sec

Title: US-09-878-781-4

Perfect score: 1715

Sequence: 1 MVVKGINGEGRIGRLAFRR.....EMSYTAQLVRLTEYFAKAK 336

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1711	99.8	336	2	A42963 glyceralddehyde-3-p
2	1617	94.3	336	2	S71350 glyceralddehyde-3-p
3	1593.5	92.9	335	2	F95235 glyceralddehyde-3-p
4	1593.5	92.9	359	2	G98099 glyceralddehyde-3-p
5	1410	82.2	336	2	F86905 glyceralddehyde-3-p
6	1338.5	78.0	337	2	F86694 hypothetical prote
7	1289.5	75.2	334	2	S34254 glyceralddehyde-3-p
8	1262	73.6	336	2	AC1382 glyceralddehyde-3-p
9	1261	73.5	336	2	AD1751 glyceralddehyde-3-p
10	1259.5	73.4	334	2	C96987 glyceralddehyde-3-p
11	1197.5	69.8	334	2	B82019 glyceralddehyde-3-p
12	1184.5	69.1	334	2	E81001 glyceralddehyde-3-p
13	1183	69.0	336	2	E89850 glyceralddehyde-3-p
14	1003	58.5	338	2	F09633 glyceralddehyde-3-p
15	993	57.9	333	2	F93881 glyceralddehyde-3-p
16	993	57.9	333	2	C85737 glyceralddehyde-3-p
17	921	53.0	337	2	S73737 glyceralddehyde-3-p
18	909.5	53.0	336	2	A43260 glyceralddehyde-3-p
19	908	52.9	337	2	C64233 glyceralddehyde-3-p
20	907	52.9	349	2	F90517 glyceralddehyde-3-p
21	886.5	52.3	335	1	DB86G glyceralddehyde-3-p
22	881.5	52.0	335	1	S12896 glyceralddehyde-3-p
23	885.5	51.6	335	1	DEB8GF glyceralddehyde-3-p
24	882.5	51.5	335	1	H84094 glyceralddehyde-3-p
25	874	51.0	333	1	DEH8GT glyceralddehyde-3-p
26	863	50.3	335	2	A70107 probable glycerald
27	843	49.2	342	2	F70391 glyceralddehyde-3-p
28	842.5	49.1	336	2	T36020 glyceralddehyde-3-p
29	831	48.5	334	2	AI0262 glyceralddehyde-3-p

30	830	48.4	331	2	F82131 glyceralddehyde-3-p
31	823	48.0	330	2	E75408 glyceralddehyde-3-p
32	822	47.9	331	2	AG0711 glyceralddehyde-3-p
33	818	47.7	331	1	DEECG3 glyceralddehyde-3-p
34	818	47.7	332	2	H90939 glyceralddehyde-3-p
35	818	47.7	332	2	D85788 glyceralddehyde-3-p
36	816	47.6	339	2	G70915 glyceralddehyde-3-p
37	815	47.5	339	2	S72763 glyceralddehyde-3-p
38	807.5	47.1	339	2	G64041 glyceralddehyde-3-p
39	799	46.6	331	1	DEUNGC glyceralddehyde-3-p
40	787.5	45.9	336	2	T10235 glyceralddehyde-3-p
41	781	45.5	337	2	S54141 glyceralddehyde-3-p
42	779.5	45.5	331	1	DETWG3 glyceralddehyde-3-p
43	779.5	45.5	337	2	S40610 glyceralddehyde-3-p
44	777.5	45.3	332	2	JC4373 glyceralddehyde-3-p
45	775	45.2	344	2	B84043 glyceralddehyde-3-p

ALIGNMENTS

RESULT 1

A42963 glyceralddehyde-3-phosphate dehydrogenase (phosphorylating) (BC 1.2.1.12) - Streptococci
N:Alternate names: plasmin receptor
C:Species: Streptococcus sp.
C:Date: 10-Mar-1994 #sequence revision 10-Mar-1994 #text_change 03-Jun-2002
C:Accession: A42963; B42963; JH0750
R:Rottenberg, R.; Broder, C.C.; Boyle, M.D.; Kain, S.J.; Schroeder, B.L.; Curries III, U. Bacteriol. 174, 5204-5210, 1992
U: Bacteriol. 174, 5204-5210, 1992
A:Title: Cloning, sequence analysis, and expression in *Escherichia coli* of a streptococcal
A:Reference number: A42963; PMID:1322883
A:Accession: A42963
A:Molecule type: DNA
A:Residues: 1-336 <LOT>
A:Experimental source: group A, strain 64/14
A:Note: sequence extracted from NCBI backbone (NCBI:110308)
A:Accession: B42963
A:Molecule type: protein
A:Residues: 2-74;161-164, 'X', 166-174;187-211, 'X', 213-217 <LOT>
R:Pancholi, V.; Fischetti, V.A.
U: Exp. Med. 176, 415-426, 1992
A:Title: A major surface protein on group A streptococci is a glyceralddehyde-3-phosphat
A:Reference number: JH0750; PMID:92364544; PMID:1500854
A:Accession: JH0750
A:Molecule type: protein
A:Residues: 2-30, 'A', 32-40 <PAN>
C:Superfamily: glyceralddehyde-3-phosphate dehydrogenase
C:Keywords: gluconeogenesis; glycolysis; homotetramer; NAD; oxidoreductase
R:152/Active site: Cys #status predicted

Query Match 99.8%; Score 1711; DB 2; Length 336;

Best Local Similarity 99.4%; Pred. No. 4.2e-116;

Matches 334; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY	1	MVVKGINGEGRIGRLAFRRIONVEGEVTRINDITDPMALHLKYDTGGRPGIVEV	60
DB	1	MVVKGINGEGRIGRLAFRRIONVEGEVTRINDITDPMALHLKYDTGGRPGIVEV	60
QY	61	KGGFEVNFIVKASRPENDIMADGVEIVLATGFEPAKKAEXKHLNAGAKKVI	120
DB	61	KGGFEVNFIVKASRPENDIMADGVEIVLATGFEPAKKAEXKHLNAGAKKVI	120
QY	121	TAPGNDVKTVPNTNHDILDTETVSGASTTNCALPMAKALHDAFGIOGKMTTTHA	180
DB	121	TAPGNDVKTVPNTNHDILDTETVSGASTTNCALPMAKALHDAFGIOGKMTTTHA	180
QY	181	YTGQDMITLDPHRRGDLRRAPAGANTVPSGAAKATIGVITPELNKIDGAAGVPPPT	240
DB	181	YTGQDMITLDPHRRGDLRRAPAGANTVPSGAAKATIGVITPELNKIDGAAGVPPPT	240
QY	241	GSFTEIVTLTKVSVDEINAMAAASNSFGYTEDPIVSSDIYGVYSGLFDATOTKVM	300
DB	241	GSFTEIVTLTKVSVDEINAMAAASNSFGYTEDPIVSSDIYGVYSGLFDATOTKVM	300

Db 241 GSVTELVTLTKNVSVCEINSAKMAKASNDSEFGTEDEPIVSSDIYVSGSLDPATQTKVM 300
QY 301 EVDGSQLYKVVSWYDNEKSYTAQVRLTEYFAKIAK 336
Db 301 EVDGSQLYKVVSWYDNEKSYTAQVRLTEYFAKIAK 336

RESULT 2

S71350

glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) - Streptococcus

C:Species: Streptococcus "equisimilis"

C:Date: 29-Jan-1998 #sequence_revision 06-Feb-1998 #text_change 03-Jun-2002

C:Accession: S71350

R:Gase, K.; Gase, A.; Schirmer, H.; Malke, H.

Eur. J. Biochem. 239, 42-51, 1996

A:Title: Cloning, sequencing and functional overexpression of the Streptococcus equisimi-

linding protein. Purification and biochemical characterization of the protein.

A:Reference number: S71350; MOID:96053564; PMID:8706717

A:Accession: S71350

A:Molecule type: DNA

A:Residues: 1-336 <GAS>

A:Cross-references: EMBL:X97788; NID:g1478268; EIDN:CA66377.1; PID:g1478269

A:Experimental source: strain H46A

C:Genetics:

A:Gene: gapC

C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase

C:Keywords: gluconogenesis; glycolysis; NAD; oxidoreductase

F:4-34:Region: beta-alpha-beta NAD nucleotide-binding fold

F:152/Active site: Cys #status predicted

Query Match 94.3%; Score 1617; DB 2; Length 336;
Best local similarity 95.5%; Pred. No. 2,5e-109;
Matches 321; Conservative 2; Mismatches 13; Indels 0; Gaps 0;

QY 1 MMYKVGINGGRIGLAFRIQNVGEVETRIINDLTPMIAHLKYDTTQGRFDGVEV 60
Db 1 MMYKVGINGGRIGLAFRIQNVGEVETRIINDLTPMIAHLKYDTTQGRFDGVEV 60
QY 61 KEGFEVNGNFKVSAERDPENIDMATDGEIVLEATGFPFAKKEAAEKHLHANGAKKVI 120
Db 61 KEGFEVNGNFKVSAERDPENIDMATDGEIVLEATGFPFAKKEAAEKHLHANGAKKVI 120
QY 121 TAPGNDVKTVEFNTNHDILDTETVTSAGSCTTNCIAPAKALQDQVGAORVPPT 180
Db 121 TAPGNDVKTVEFNTNHDILDTETVTSAGSCTTNCIAPAKALQDQVGAORVPPT 180
QY 181 YTGDDMTLQEPHGGDLRRARAGANIVPNSGTAAKATGLVPELNGKLDGAAGRPVPT 240
Db 181 YTGDDMTLQEPHGGDLRRARAGANIVPNSGTAAKATGLVPELNGKLDGAAGRPVPT 240
QY 241 GSVTELVTLTKNVSVCEINSAKMAKASNDSEFGTEDEPIVSSDIYVSGSLDPATQTKVM 300
Db 241 GSVTELVTLTKNVSVCEINSAKMAKASNDSEFGTEDEPIVSSDIYVSGSLDPATQTKVM 300
QY 301 EVDGSQLYKVVSWYDNEKSYTAQVRLTEYFAKIAK 336
Db 301 EVDGSQLYKVVSWYDNEKSYTAQVRLTEYFAKIAK 336

RESULT 3

F95235

glyceraldehyde-3-phosphate dehydrogenase [imported] - Streptococcus pneumoniae (strain T

C:Species: Streptococcus pneumoniae

C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 24-Aug-2001

C:Accession: F95235

R:Rettelid, H.; Neilson, K.E.; Paulsen, I.T.; Eiserich, J.A.; Read, T.D.; Peterson, S.; Heid

cn, U.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holzapfel,

son, T.; Hickey, E.K.; Holt, I.E.

Science 293, 498-506, 2001

A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,

A:Title: Complete Genome Sequence of a Virulent Isolate of Streptococcus pneumoniae.

A:Reference number: A95000; MUID:21357209; PMID:11469396

A:Accession: F95235

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-335 <GUR>
A:Cross-references: GB:A800672; PIDN:AAK76079.1; PID:g14973522; GSEDB:GN00164; TIGR:
A:Experimental source: strain TIGR4
C:Genetics:
A:Gene: SP2012
C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase

Query Match 92.9%; Score 1593.5; DB 2; Length 335;
Best local similarity 92.6%; Pred. No. 1.3e-107;
Matches 311; Conservative 12; Mismatches 12; Indels 1; Gaps 1;

QY 1 MMYKVGINGGRIGLAFRIQNVGEVETRIINDLTPMIAHLKYDTTQGRFDGVEV 60
Db 1 MMYKVGINGGRIGLAFRIQNVGEVETRIINDLTPMIAHLKYDTTQGRFDGVEV 60
QY 61 KEGFEVNGNFKVSAERDPENIDMATDGEIVLEATGFPFAKKEAAEKHLHANGAKKVI 120
Db 61 KEGFEVNGNFKVSAERDPENIDMATDGEIVLEATGFPFAKKEAAEKHLHANGAKKVI 119
QY 121 TAPGNDVKTVEFNTNHDILDTETVTSAGSCTTNCIAPAKALQDQVGAORVPPT 180
Db 121 TAPGNDVKTVEFNTNHDILDTETVTSAGSCTTNCIAPAKALQDQVGAORVPPT 179
QY 181 YTGDDMTLQEPHGGDLRRARAGANIVPNSGTAAKATGLVPELNGKLDGAAGRPVPT 240
Db 181 YTGDDMTLQEPHGGDLRRARAGANIVPNSGTAAKATGLVPELNGKLDGAAGRPVPT 239
QY 241 GSVTELVTLTKNVSVCEINSAKMAKASNDSEFGTEDEPIVSSDIYVSGSLDPATQTKVM 300
Db 241 GSVTELVTLTKNVSVCEINSAKMAKASNDSEFGTEDEPIVSSDIYVSGSLDPATQTKVM 299
QY 301 EVDGSQLYKVVSWYDNEKSYTAQVRLTEYFAKIAK 336
Db 301 EVDGSQLYKVVSWYDNEKSYTAQVRLTEYFAKIAK 335

RESULT 4

G98099

glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) [imported] -

C:Species: Streptococcus pneumoniae

C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 03-Jun-2002

C:Accession: G98099

R:Joshkins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; Dehoff, B.S.;

e, R.; LeBlanc, D.J.; Lee, L.N.; Iefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.;

J. Bacteriol. 183, 5709-5717, 2001

A:Authors: Yang, Y.; Young-Belland, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.;

A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.

A:Reference number: A97872; MUID:21429245; PMID:11544234

A:Accession: G98099

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-359 <GUR>

A:Cross-references: GB:A8007317; PIDN:AAH00628.1; PID:g15459513; GSEDB:GN00174

C:Genetics:

A:Gene: gapA

C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase

C:Keywords: oxidoreductase

Query Match 92.9%; Score 1593.5; DB 2; Length 359;
Best local similarity 92.6%; Pred. No. 1.4e-107;
Matches 311; Conservative 12; Mismatches 12; Indels 1; Gaps 1;

QY 1 MMYKVGINGGRIGLAFRIQNVGEVETRIINDLTPMIAHLKYDTTQGRFDGVEV 60
Db 1 MMYKVGINGGRIGLAFRIQNVGEVETRIINDLTPMIAHLKYDTTQGRFDGVEV 60
QY 25 MMYKVGINGGRIGLAFRIQNVGEVETRIINDLTPMIAHLKYDTTQGRFDGVEV 84
Db 25 MMYKVGINGGRIGLAFRIQNVGEVETRIINDLTPMIAHLKYDTTQGRFDGVEV 84
QY 61 KEGFEVNGNFKVSAERDPENIDMATDGEIVLEATGFPFAKKEAAEKHLHANGAKKVI 120
Db 61 KEGFEVNGNFKVSAERDPENIDMATDGEIVLEATGFPFAKKEAAEKHLHANGAKKVI 120
QY 85 KEGFEVNGNFKVSAERDPENIDMATDGEIVLEATGFPFAKKEAAEKHLHANGAKKVI 143
Db 85 KEGFEVNGNFKVSAERDPENIDMATDGEIVLEATGFPFAKKEAAEKHLHANGAKKVI 143

QY 121 TAPGNDVKTVPVNTNHDILDTETVTSAGSCTTNCIAFMAYALHDAFGIOKGLMTTHA 180
DB 144 TAPGNDVKTVPVNTNHDILDTETVTSAGSCTTNCIAFMAYALHDAFGIOKGLMTTHA 203
QY 181 YTDQMTLDGPHRGDIPRRARAAANIVPNSGTAAKATGIVPELNGKLDGAAGRPVPT 240
DB 204 YTDQMTLDGPHRGDIPRRARAAANIVPNSGTAAKATGIVPELNGKLDGAAGRPVPT 263
QY 241 GSVTELVVTLDKNVSDVEINAAKASNDSPGTEDPIYASDVGVSGLPFAQTQKYM 300
DB 264 GSVTELVVTLDKNVSDVEINAAKASNDSPGTEDPIYASDVGVSGLPFAQTQKYM 323
QY 301 EVDGSQLVXVSWYDNEMSYTAQVRLTEYPAKIAK 336
DB 324 DVGKQVLVXVSWYDNEMSYTAQVRLTEYPAKIAK 359

RESULT 5

F86905

hypothetical protein gapB [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C:Species: Lactococcus lactis subsp. lactis
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C:Accession: F86905

R:Botolin, A.; Munker, P.; Manger, S.; Jallou, O.; Malarme, K.; Weissenbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s
A:Reference number: A86625; MUID:21235186; PMID:11337471
A:Accession: F86905

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-336 <STO>

A:Cross-references: GB:A8005176; PID:g12725315; PIDN:AAK0634.1; GSPDB:GN00146

A:Experimental source: strain IL1403

A:Genetics:

A:Gene: gapB

C:Superfamily: glyceroldehyde-3-phosphate dehydrogenase

Query Match 82.2%; Score 1410; DB 2; Length 336;
Best Local Similarity 81.2%; Pred. No. 2.2e-94;
Matches 273; Conservative 22; Mismatches 41; Indels 0; Gaps 0;

QY 1 MVAKVGINGRIGRLAFRRIQVGEVETRIINDLTDPMALHLKDYTQGRFPGTVEV 60
DB 1 MVAKVGINGRIGRLAFRRIQVGEVETRIINDLTDPMALHLKDYTQGRFPGTVEV 60
QY 61 KEGFEVNGKFIKVAERDPENIDMADGVEIVLAEATGFAKKEAEXKHLHANGAKKVI 120
DB 61 KEGFEVNGKFIKVAERDPENIDMADGVEIVLAEATGFAKKEAEXKHLHANGAKKVI 120
QY 121 TAPGNDVKTVPVNTNHDILDTETVTSAGSCTTNCIAFMAYALHDAFGIOKGLMTTHA 180
DB 121 TAPGNDVKTVPVNTNHDILDTETVTSAGSCTTNCIAFMAYALHDAFGIOKGLMTTHA 180
QY 181 YTDQMTLDGPHRGDIPRRARAAANIVPNSGTAAKATGIVPELNGKLDGAAGRPVPT 240
DB 181 YTDQMTLDGPHRGDIPRRARAAANIVPNSGTAAKATGIVPELNGKLDGAAGRPVPT 240
QY 241 GSVTELVVTLDKNVSDVEINAAKASNDSPGTEDPIYASDVGVSGLPFAQTQKYM 300
DB 241 GSVTELVVTLDKNVSDVEINAAKASNDSPGTEDPIYASDVGVSGLPFAQTQKYM 300
QY 301 EVDGSQLVXVSWYDNEMSYTAQVRLTEYPAKIAK 336
DB 301 EVDGSQLVXVSWYDNEMSYTAQVRLTEYPAKIAK 336

RESULT 6

G86694

hypothetical protein gapB [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C:Species: Lactococcus lactis subsp. lactis
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C:Accession: G86694

R:Botolin, A.; Munker, P.; Manger, S.; Jallou, O.; Malarme, K.; Weissenbach, J.; Ehrlich

Genome Res. 11, 731-753, 2001
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis
A:Reference number: A86625; MUID:21235186; PMID:11337471
A:Accession: G86694
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-337 <STO>
A:Cross-references: GB:A8005176; PID:g12723446; PIDN:AAK04657.1; GSPDB:GN00146
A:Experimental source: strain IL1403
A:Genetics:

A:Gene: gapB

C:Superfamily: glyceroldehyde-3-phosphate dehydrogenase

Query Match 78.0%; Score 1338.5; DB 2; Length 337;
Best Local Similarity 77.2%; Pred. No. 3.2e-89;
Matches 260; Conservative 31; Mismatches 45; Indels 1; Gaps 1;

QY 1 MVAKVGINGRIGRLAFRRIQVGEVETRIINDLTDPMALHLKDYTQGRFPGTVEV 60
DB 1 MVAKVGINGRIGRLAFRRIQVGEVETRIINDLTDPMALHLKDYTQGRFPGTVEV 60
QY 61 KEGFEVNGKFIKVAERDPENIDMADGVEIVLAEATGFAKKEAEXKHLHANGAKKVI 120
DB 61 KEGFEVNGKFIKVAERDPENIDMADGVEIVLAEATGFAKKEAEXKHLHANGAKKVI 120
QY 121 TAPGNDVKTVPVNTNHDILDTETVTSAGSCTTNCIAFMAYALHDAFGIOKGLMTTHA 180
DB 121 TAPGNDVKTVPVNTNHDILDTETVTSAGSCTTNCIAFMAYALHDAFGIOKGLMTTHA 180
QY 181 YTDQMTLDGPHRGDIPRRARAAANIVPNSGTAAKATGIVPELNGKLDGAAGRPVPT 240
DB 181 YTDQMTLDGPHRGDIPRRARAAANIVPNSGTAAKATGIVPELNGKLDGAAGRPVPT 240
QY 241 GSVTELVVTLDKNVSDVEINAAKASNDSPGTEDPIYASDVGVSGLPFAQTQKYM 300
DB 241 GSVTELVVTLDKNVSDVEINAAKASNDSPGTEDPIYASDVGVSGLPFAQTQKYM 300
QY 301 EVDGSQLVXVSWYDNEMSYTAQVRLTEYPAKIAK 336
DB 301 EVDGSQLVXVSWYDNEMSYTAQVRLTEYPAKIAK 337

RESULT 7

S34254

glyceroldehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) - Clostridium
C:Species: Clostridium pasteurianum
C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 03-Jun-2002
C:Accession: S34254

R:Oster, T.; Assoheli, O.; Scherzer, S.; Brantant, G.; Brantant, C.

submitted to the EMBL data library, May 1993

A:Description: Nucleotide sequence of the glyceroldehyde-3-phosphate dehydrogenase from

A:Reference number: S34254

A:Accession: S34254

A:Molecule type: DNA

A:Residues: 1-334 <OST>

A:Cross-references: EMBL:X72219; NID:g311923; PIDN:CAAS1020.1; PID:g311924

C:Superfamily: glyceroldehyde-3-phosphate dehydrogenase

C:Keywords: oxidoreductase

Query Match 75.2%; Score 1289.5; DB 2; Length 334;
Best Local Similarity 74.3%; Pred. No. 1.1e-85;
Matches 249; Conservative 33; Mismatches 52; Indels 1; Gaps 1;

QY 2 VVAKVGINGRIGRLAFRRIQVGEVETRIINDLTDPMALHLKDYTQGRFPGTVEV 61
DB 1 MTKVAINGRIGRLAFRRILIEVPGLEVAINDLTDPMALHLKDYTQGRFPGTVEV 60
QY 62 EGFEVNGKFIKVAERDPENIDMADGVEIVLAEATGFAKKEAEXKHLHANGAKKVI 121
DB 62 EGFEVNGKFIKVAERDPENIDMADGVEIVLAEATGFAKKEAEXKHLHANGAKKVI 121
QY 121 TAPGNDVKTVPVNTNHDILDTETVTSAGSCTTNCIAFMAYALHDAFGIOKGLMTTHA 180
DB 121 TAPGNDVKTVPVNTNHDILDTETVTSAGSCTTNCIAFMAYALHDAFGIOKGLMTTHA 180
QY 181 YTDQMTLDGPHRGDIPRRARAAANIVPNSGTAAKATGIVPELNGKLDGAAGRPVPT 240
DB 181 YTDQMTLDGPHRGDIPRRARAAANIVPNSGTAAKATGIVPELNGKLDGAAGRPVPT 240
QY 241 GSVTELVVTLDKNVSDVEINAAKASNDSPGTEDPIYASDVGVSGLPFAQTQKYM 300
DB 241 GSVTELVVTLDKNVSDVEINAAKASNDSPGTEDPIYASDVGVSGLPFAQTQKYM 300
QY 301 EVDGSQLVXVSWYDNEMSYTAQVRLTEYPAKIAK 336
DB 301 EVDGSQLVXVSWYDNEMSYTAQVRLTEYPAKIAK 337

Db 120 APAGNDKTIIVFWNNEDLDGTETIVISGASCTTNCIAAPAKVINDKEGKPMTHIAY 179
Cy 182 TSDGMILGPHRGDGPBARAGANIVPNTSGAAYAGLVIPELNGKLDGAQRVPYTG 241
Db 180 TNDQNTLDGPHRGDGPBARAGANIVPNTSGAAYAGLVIPELNGKLDGAQRVPYTG 239
Cy 242 SVTELVTYLDKNVSVDEINAAKASNDSPGYEDPIYSDIVSGYSLPDTOTKME 301
Db 240 SVTELISVLKKNVTVEIINAAKAAESGTYEDEIVSADVIGISLFPATLTIKVD 299
Cy 302 VDGSLYKVVSVTNDENMSTYQVLTTEYFAKIAK 336
Db 300 VDGSLYKVVSVTNDENMSTYQVLTTEYFAKIAK 334

RESULT 8

AC1382
glyceraldehyde 3-phosphate dehydrogenase homolog gap [imported] - Listeria monocytogenes
C:Species: Listeria monocytogenes
C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C:Accession: AC1382
R:Glaeser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Jones, L.M.; Kars, U.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, E.; Science 294, 849-852, 2001
A:Authors: Kref, U.; Kuhn, M.; Kunst, F.; Kurapkac, G.; Madueno, E.; Maitouram, A.; Ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlant, A.; Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AC1382
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-336 <GLA>
A:Cross-references: GB:NC_003210; PIDD:CAD00537.1; PIDD:g14151947; GSPDB:GN00177
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: gap
A:Superfamily: glyceraldehyde-3-phosphate dehydrogenase

Query Match 73.6%; Score 1262; DB 2; Length 336;
Best Local Similarity 73.3%; Pred. No. 1.1e-83;
Matches 247; Conservative 33; Mismatches 55; Indels 2; Gaps 2;

Cy 1 MVTAKVINGFGRIGRLAFRIQNVGVEVTRINDLTPNMLAHLKYDTTGGFPGTVEV 60
Db 1 MVTAKVINGFGRIGRLAFRIQNVGVEVTRINDLTPNMLAHLKYDTTGGFPGTVEV 60
Cy 61 KEGGFVNGFNIKVSARDPENIDMADGVEIVLEATGFPKAKAAKHANAKKVI 120
Db 61 HDGFVNGFNIKVSARDPENIDMADGVEIVLEATGFPKAKAAKHANAKKVI 119
Cy 121 TAPGNDVTVTFVFNHDLIDGTETIVISGASCTTNCIAAPAKVINDKEGKPMTHIAY 180
Db 120 SAPATGDKTIVVNNHETLDGTETIVISGASCTTNCIAAPAKVINDKEGKPMTHIAY 179
Cy 181 YTGDMILDGPHRGDGPBARAGANIVPNTSGAAYAGLVIPELNGKLDGAQRVPYTG 240
Db 180 YTGDMILDGPHRGDGPBARAGANIVPNTSGAAYAGLVIPELNGKLDGAQRVPYTG 239
Cy 241 GSVELVTYLDKNVSVDEINAAKASNDSPGYEDPIYSDIVSGYSLPDTOTKME 301
Db 240 GSVELVTYLDKNVSVDEINAAKASNDSPGYEDPIYSDIVSGYSLPDTOTKME 299
Cy 300 MEVDGSLYKVVSVTNDENMSTYQVLTTEYFAKIAK 336
Db 300 MEVDGSLYKVVSVTNDENMSTYQVLTTEYFAKIAK 336

RESULT 9

AD1751
glyceraldehyde 3-phosphate dehydrogenase homolog gap [imported] - Listeria innocua (strain C)
C:Species: Listeria innocua
C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001

C:Accession: AD1751
R:Glaeser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, E.; Jones, L.M.; Kars, U.; Science 294, 849-852, 2001
A:Authors: Kref, U.; Kuhn, M.; Kunst, F.; Kurapkac, G.; Madueno, E.; Maitouram, A.; Ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlant, A.; Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AD1751
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-336 <GLA>
A:Cross-references: GB:AL592022; PIDD:CAAC57780.1; PIDD:g1415075; GSPDB:GN00178
A:Experimental source: strain Clp11262
C:Genetics:
A:Gene: gap
A:Superfamily: glyceraldehyde-3-phosphate dehydrogenase

Query Match 73.5%; Score 1261; DB 2; Length 336;
Best Local Similarity 73.3%; Pred. No. 1.1e-83;
Matches 247; Conservative 32; Mismatches 56; Indels 2; Gaps 2;

Cy 1 MVTAKVINGFGRIGRLAFRIQNVGVEVTRINDLTPNMLAHLKYDTTGGFPGTVEV 60
Db 1 MVTAKVINGFGRIGRLAFRIQNVGVEVTRINDLTPNMLAHLKYDTTGGFPGTVEV 60
Cy 61 KEGGFVNGFNIKVSARDPENIDMADGVEIVLEATGFPKAKAAKHANAKKVI 120
Db 61 HDGFVNGFNIKVSARDPENIDMADGVEIVLEATGFPKAKAAKHANAKKVI 119
Cy 121 TAPGNDVTVTFVFNHDLIDGTETIVISGASCTTNCIAAPAKVINDKEGKPMTHIAY 180
Db 120 SAPATGDKTIVVNNHETLDGTETIVISGASCTTNCIAAPAKVINDKEGKPMTHIAY 179
Cy 181 YTGDMILDGPHRGDGPBARAGANIVPNTSGAAYAGLVIPELNGKLDGAQRVPYTG 240
Db 180 YTGDMILDGPHRGDGPBARAGANIVPNTSGAAYAGLVIPELNGKLDGAQRVPYTG 239
Cy 241 GSVELVTYLDKNVSVDEINAAKASNDSPGYEDPIYSDIVSGYSLPDTOTKME 301
Db 240 GSVELVTYLDKNVSVDEINAAKASNDSPGYEDPIYSDIVSGYSLPDTOTKME 299
Cy 300 MEVDGSLYKVVSVTNDENMSTYQVLTTEYFAKIAK 336
Db 300 MEVDGSLYKVVSVTNDENMSTYQVLTTEYFAKIAK 336

RESULT 10

C96987
glyceraldehyde 3-phosphate dehydrogenase, gene gapC [imported] - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum
C>Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 30-Sep-2001
C:Accession: C96987
R:Rolling, U.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Le, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.; J. Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clostridium acetobutylicum ATCC824
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: C96987
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-334 <KOR>
A:Cross-references: GB:AE001437; PIDD:AAK78686.1; PIDD:g15023589; GSPDB:GN00168
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
A:Gene: CAC0709
A:Superfamily: glyceraldehyde-3-phosphate dehydrogenase

Query Match 73.4%; Score 1259.5; DB 2; Length 334;
Best Local Similarity 71.9%; Pred. No. 1.6e-83;
Matches 241; Conservative 40; Mismatches 53; Indels 1; Gaps 1;

QY 2 VYKVGINGRIGRLAARRIQWVSEVTRINDLTPNMLHLKYDTTQGRPDGVYEV 61
1 MAKIANGRIGRLAARRIQWVSEVTRINDLTPNMLHLKYDTTQGRPDGVYEV 60
QY 62 EGGEFVNGKFIKVAEADPENIDMTDGVETVLAAGTFPAKKEAAEKHLHANGAKKVI 121
Db 61 EGAFVNGKFIKVAEADPENIDMTDGVETVLAAGTFPAKKEAAEKHLHANGAKKVI 119
QY 122 AAGNDKTVVNTNHDIDGTETVLSGASCTTNCLAPMAKALHARIGOKGMTTHAY 181
Db 120 AAGNDKTVVNTNHDIDGTETVLSGASCTTNCLAPMAKALHARIGOKGMTTHAY 179
QY 182 TGDQMLDGPGRGDLRRARAGANIVNSGAAKAGIVIPELAKKLDAAGRVPEPTG 241
Db 180 TNDQMLDGPGRGDLRRARAGANIVNSGAAKAGIVIPELAKKLDAAGRVPEPTG 239
QY 242 SYTELAVTLTKVNSVDENINAAKKAASNDSPGYTDPVSSDIVGVSGLPDAQTCKWE 301
Db 240 SYTELAVTLTKVNSVDENINAAKKAASNDSPGYTDPVSSDIVGVSGLPDAQTCKWE 299
QY 302 VDSGLVAVSWYNENMSYTAQVRLTEYFKAK 336
Db 300 VDSGLVAVSWYNENMSYTAQVRLTEYFKAK 334

RESULT 11
B82019
glyceralddehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) C:MM2046 [imp
C:Species: Neisseria meningitidis
C>Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 03-Jun-2002
C:Accession: B82019
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
J.; Holtroyd, S.; Jørgensen, K.; Leather, S.; Mouton, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A>Title: Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491.
A:Reference number: A81775; MUID:2022556; PMID:10761919
A:Accession: B82019
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-334 <PAR>
A:Cross-references: GB:A162752; GB:A157959; NID:g7378778; PIND:CA83554.1; PID:g737902
A:Experimental source: serogroup A, strain Z2491
A:Genetic: C:Genetics
A:Gene: MM2046
C:Superfamily: glyceralddehyde-3-phosphate dehydrogenase
C:Keywords: oxidoreductase

Query Match 69.8%; Score 1197.5; DB 2; Length 334;
Best local similarity 70.2%; Pred. No. 4.8e-79;
Matches 233; Conservative 33; Mismatches 65; Indels 1; Gaps 1;

QY 1 MYYKVGINGRIGRLAARRIQWVSEVTRINDLTPNMLHLKYDTTQGRPDGVYEV 60
Db 1 MSIKVANGRIGRLAARRIQWVSEVTRINDLTPNMLHLKYDTTQGRPDGVYEV 60
QY 61 KEGGEFVNGKFIKVAEADPENIDMTDGVETVLAAGTFPAKKEAAEKHLHANGAKKVI 120
Db 61 KDAIVNGKFIKVAEADPENIDMTDGVETVLAAGTFPAKKEAAEKHLHANGAKKVI 119
QY 121 TAPGNDVKTVPNTNHDIDGTETVLSGASCTTNCLAPMAKALHARIGOKGMTTHA 180
Db 120 TAPGNDVKTVPNTNHDIDGTETVLSGASCTTNCLAPMAKALHARIGOKGMTTHA 179
QY 181 YTGQMLDGPGRGDLRRARAGANIVNSGAAKAGIVIPELAKKLDAAGRVPEPTG 240
Db 180 YTGQMLDGPGRGDLRRARAGANIVNSGAAKAGIVIPELAKKLDAAGRVPEPTG 239
QY 241 GSYTELAVTLTKVNSVDENINAAKKAASNDSPGYTDPVSSDIVGVSGLPDAQTCKW 300
Db 240 GSYTELAVTLTKVNSVDENINAAKKAASNDSPGYTDPVSSDIVGVSGLPDAQTCKW 299
QY 301 EVDGSLVAVSWYNENMSYTAQVRLTEYFA 332
Db 301 EVDGSLVAVSWYNENMSYTAQVRLTEYFA 332

Db 300 TVGKQLVKTVAWYNENMSYTAQVRLTEYFA 331

RESULT 12
B81001
glyceralddehyde 3-phosphate dehydrogenase NMB2159 [imported] - *Neisseria meningitidis* (
C:Species: *Neisseria meningitidis*
C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C:Accession: B81001
R:Teichlin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.
Li, H.; Qin, H.; Vamathevan, D.; Gill, J.; Scarlato, V.; Mastignani, V.; Piazza, N.
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.;
A>Title: Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58.
A:Reference number: A81000; MUID:20175755; PMID:10710307
A:Accession: B81001
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-334 <TEXT>
A:Cross-references: GB:A8002563; GB:A8002096; NID:g7227405; PIND:AAF2467.1; PID:g7227
A:Experimental source: serogroup B, strain MC58
C:Genetic: C:Genetics
A:Gene: NMB2159
C:Superfamily: glyceralddehyde-3-phosphate dehydrogenase

Query Match 69.1%; Score 1184.5; DB 2; Length 334;
Best local similarity 69.3%; Pred. No. 4.1e-78;
Matches 230; Conservative 35; Mismatches 66; Indels 1; Gaps 1;

QY 1 MYYKVGINGRIGRLAARRIQWVSEVTRINDLTPNMLHLKYDTTQGRPDGVYEV 60
Db 1 MSIKVANGRIGRLAARRIQWVSEVTRINDLTPNMLHLKYDTTQGRPDGVYEV 60
QY 61 KEGGEFVNGKFIKVAEADPENIDMTDGVETVLAAGTFPAKKEAAEKHLHANGAKKVI 120
Db 61 KDAIVNGKFIKVAEADPENIDMTDGVETVLAAGTFPAKKEAAEKHLHANGAKKVI 119
QY 121 TAPGNDVKTVPNTNHDIDGTETVLSGASCTTNCLAPMAKALHARIGOKGMTTHA 180
Db 120 TAPGNDVKTVPNTNHDIDGTETVLSGASCTTNCLAPMAKALHARIGOKGMTTHA 179
QY 181 YTGQMLDGPGRGDLRRARAGANIVNSGAAKAGIVIPELAKKLDAAGRVPEPTG 240
Db 180 YTGQMLDGPGRGDLRRARAGANIVNSGAAKAGIVIPELAKKLDAAGRVPEPTG 239
QY 241 GSYTELAVTLTKVNSVDENINAAKKAASNDSPGYTDPVSSDIVGVSGLPDAQTCKW 300
Db 240 GSYTELAVTLTKVNSVDENINAAKKAASNDSPGYTDPVSSDIVGVSGLPDAQTCKW 299
QY 301 EVDGSLVAVSWYNENMSYTAQVRLTEYFA 332
Db 300 TVGKQLVKTVAWYNENMSYTAQVRLTEYFA 331

RESULT 13
B89850
glyceralddehyde-3-phosphate dehydrogenase [imported] - *Staphylococcus aureus* (strain N3;
C:Species: *Staphylococcus aureus*
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: B89850
R:Kurita, M.; Ohta, T.; Uchiyama, T.; Baba, T.; Yuzawa, H.; Kobayashi, T.; Cui, L.; Ogi
ma, A.; Mizutani-Uti, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayaishi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A>Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: B89850
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-336 <KUR>
A:Cross-references: GB:BA000018; PID:g13700663; PIND:BA841960.1; GSPDB:GN00149
A:Experimental source: strain N315

C:Genetics:
A:Gene: gap
C:Superfamily: glyceralddehyde-3-phosphate dehydrogenase

Query Match 69.0%; Score 1183; DB 2; Length 336;
Best Local Similarity 68.2%; Pred. No. 5.3e-78;
Matches 230; Conservative 43; Mismatches 62; Indels 2; Gaps 2;

QY 1 MYYKVGINGFGRIGRLAFRIQNVGAVTRINDLDPNMAHLKDTTQGRGDTVEV 60
DB 1 MAAKVAINGFGRIGRLAFRIQNVGAVTRINDLDPNMAHLKDTTQGRGDTVEV 60
QY 61 KEGGEPVNGNFKYSAARDPENIDATGVEVLEAFNGFAKXBAEKHLNAGAKVYI 120
DB 61 VDSGFRVNGKESFSPDSKLPKMDINDVETGCTTQKGAQHLEA-GAKVYI 119
QY 121 TAFGANDVATVFTNHDILDTETVISAQCTTNCIAPAKALHDAFGIOKMTTTHA 180
DB 120 SAPAGDILKTIVFTNHDILDTETVISAQCTTNCIAPAKALHDAFGIOKMTTTHA 179
QY 181 YTGDDNILDGPHRGDILRRARAGANIVPNTGAAKAGLVPELNGKLDGAORVPVPT 240
DB 180 YTGDDNLTADPHRGDILRRARAGANIVPNTGAAKAGLVPELNGKLDGAORVPVPT 239
QY 241 GSVTEAVYTLDK-NVSYDEINAAKASNDSPYVSDIYVSGSLFDPATQTKV 299
DB 240 GSITELTYVLRKQDVTEQVNMKNASNEFSGTDEIVSSDVGMVTSFLDPATQTKV 299
QY 300 MEVDSSQLVYVSVYDNEMSYTAOLVETLEFPAKAK 336
DB 300 MSVGRQLYKVAAYDNEMSYTAOLVETLEFPAKAK 336

RESULT 14

T09633

glyceralddehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) - Lactobacillus
C:Species: Lactobacillus delbrueckii

C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 03-Jun-2002

C:Accession: T09633

R:Barany, F.; Delacorte, F.; Garrel, J.R.

Microbiology 144, 905-914, 1998

A:Title: An operon coding for three glycolytic enzymes in Lactobacillus delbrueckii subs

A:Reference number: Z16788; MUID:98240227; PMID:9579064

A:Accession: T09633

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-338 <EPR>

A:Cross-references: EMBL:AJ000339; NID:G2624189; PIDN:CAA04014.1; PID:G2624191

A:Experimental source: subsp. bulgaricus, strain B107

C:Genetics:

A:Gene: gap

C:Superfamily: glyceralddehyde-3-phosphate dehydrogenase

C:Keywords: gluconeogenesis; glycolysis; oxidoreductase

Query Match 58.5%; Score 1003; DB 2; Length 338;

Best Local Similarity 58.7%; Pred. No. 5.2e-65;

Matches 199; Conservative 45; Mismatches 89; Indels 6; Gaps 3;

QY 1 MYYKVGINGFGRIGRLAFRIQNVGAVTRINDLDPNMAHLKDTTQGRGDTVEV 56
DB 1 MAAKVAINGFGRIGRLAFRIQNVGAVTRINDLDPNMAHLKDTTQGRGDTVEV 60
QY 57 TVEWKEGFEVNGNFKYSAARDPENIDATGVEVLEAFNGFAKXBAEKHLNAG 115
DB 61 EVSALTEDSLVWDGKKEVYAEPOANIIPVANKGIDVLECTGFTTSKASQAHLEA-GA 119
QY 116 KXVITAPAGNDVKTIVFTNHDILDTETVISAQCTTNCIAPAKALHDAFGIOK 175
DB 120 KXVITAPAGNDVKTIVFTNHDILDTETVISAQCTTNCIAPAKALHDAFGIOK 179
QY 176 TTHAYTATQKVLDPGRGDLRRARAGANIVPNTGAAKAGLVPELNGKLDGAOR 235

DB 180 TTHAYTATQKVLDPGRGDLRRARAGANIVPNTGAAKAGLVPELNGKLDGAOR 239

QY 236 VPPGSLVETLVTLDDKRVSVDEINAAKASNDSPYVSDIYVSGSLFDPAT 295

DB 240 VPPGSLVETLVTLDDKRVSVDEINAAKASNDSPYVSDIYVSGSLFDPAT 299

QY 296 QTKYMEVDSSQLVYVSVYDNEMSYTAOLVETLEFPAKAK 334

DB 300 QTKYMEVDSSQLVYVSVYDNEMSYTAOLVETLEFPAKAK 338

RESULT 15

F90881

glyceralddehyde-3-phosphate dehydrogenase C [imported] - Escherichia coli (strain 0157

C:Species: Escherichia coli

C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001

C:Accession: F90881

R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.

Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and g

A:Reference number: A99629; MUID:21156231; PMID:11258796

A:Accession: F90881

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-333 <HAY>

A:Cross-references: GB:BA000007; PIDN:BA035445.1; PID:G13361488; GSPDB:GN00154

A:Experimental source: strain 0157:H7, substrain RIMD 0509952

C:Genetics:

A:Gene: GCS0222

C:Superfamily: glyceralddehyde-3-phosphate dehydrogenase

Query Match 57.9%; Score 993; DB 2; Length 333;

Best Local Similarity 59.3%; Pred. No. 2.7e-64;

Matches 198; Conservative 54; Mismatches 76; Indels 6; Gaps 5;

QY 4 KXVINGFGRIGRLAFRIQNVGAVTRINDLDPNMAHLKDTTQGRGDTVEV 62

DB 3 KXVINGFGRIGRLAFRIQNVGAVTRINDLDPNMAHLKDTTQGRGDTVEV 62

QY 63 GGFVNGNFKYSAARDPENIDATGVEVLEAFNGFAKXBAEKHLNAGAKVYI 122

DB 63 DSVLVDKSLVYVSVYDNEMSYTAOLVETLEFPAKAK 121

QY 123 TAFGANDVATVFTNHDILDTETVISAQCTTNCIAPAKALHDAFGIOKMTTTHA 182

DB 122 TAFGANDVATVFTNHDILDTETVISAQCTTNCIAPAKALHDAFGIOKMTTTHA 180

QY 183 GDQMLDGPFRGDLRRARAGANIVPNTGAAKAGLVPELNGKLDGAORVPVPT 242

DB 181 GDQMLDGPFRGDLRRARAGANIVPNTGAAKAGLVPELNGKLDGAORVPVPT 239

QY 243 VTELVTLDGKRVSVDEINAAKASNDSPYVSDIYVSGSLFDPATQTKV 300

DB 240 VTELVTLDGKRVSVDEINAAKASNDSPYVSDIYVSGSLFDPATQTKV 299

QY 301 EYDSSQLVYVSVYDNEMSYTAOLVETLEFPAKAK 334

DB 300 EYDSSQLVYVSVYDNEMSYTAOLVETLEFPAKAK 333

Search completed: May 5, 2004, 14:21:42
Job time: 21 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 14:15:03 ; Search time 17 Seconds

(without alignments)
1029.151 Million cell updates/sec

Title: US-09-878-781-4

Perfect score: 1715

Sequence: 1 MVVXVINGRIGRIARFR.....EMSYTAQWRTLEFVAK 336

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_42*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1709	99.7	335	1	G3P_STREY
2	1705	99.4	335	1	G3P_STRE3
3	1612	94.0	335	1	G3P_STRE0
4	1338.5	78.0	337	1	G3P_IACIA
5	1289.5	75.2	334	1	G3P_CLOPA
6	1259.5	73.4	334	1	G3P_CLOPA
7	1190	69.4	336	1	G3P_STRE
8	1183	69.0	336	1	G3P_STPAK
9	1003	58.5	338	1	G3P_LACDE
10	993	57.9	333	1	G3P_ECO57
11	992	57.8	333	1	G3P_ECOLI
12	921	53.7	337	1	G3P_MYCPN
13	908.5	53.0	334	1	G3P_CORGL
14	908	52.9	337	1	G3P_MYCGB
15	891.5	52.0	334	1	G3P_BACST
16	886.5	51.7	334	1	G3P_BACME
17	880.5	51.3	334	1	G3P_BACST
18	873	50.9	332	1	G3P_THEMA
19	863	50.3	335	1	G3P_BORBU
20	843	49.2	342	1	G3P_AQUAE
21	842.5	49.1	336	1	G3P_STRCO
22	817	47.6	330	1	G3P_SALTY
23	816	47.6	339	1	G3P_MYCTU
24	815	47.5	339	1	G3P_MYCLE
25	813	47.4	330	1	G3P_ECOLI
26	807.5	47.1	339	1	G3P_HACTN
27	801	46.7	339	1	G3P_MYCAN
28	794.5	46.3	337	1	G3P_RHIRA
29	794	46.3	330	1	G3P_TRYBB
30	787.5	45.9	336	1	G3P_SCHPO
31	784.5	45.7	332	1	G3P_RALSO
32	783.5	45.7	337	1	G3P_MOMAN
33	779.5	45.5	331	1	G3P_THENQ

34	779.5	45.5	337	1	G3P_CLAPU	O00584 claviceps p
35	777.5	45.3	332	1	G3P_STRAU	O59800 streptomyc
36	775	45.2	336	1	G3P_STRE3	P00505 streptocyst
37	772	45.0	333	1	G3P_STRE	P54226 streptomyc
38	772	45.0	337	1	G3P2_AANAS	P58554 andbiana sp
39	771.5	45.0	337	1	G3P_COCHR	P29457 cochlilobu
40	771	45.0	330	1	G3P2_LEISM	O01558 leishmania
41	770.5	44.9	335	1	G3P2_SCHPO	P28844 schizosach
42	770.5	44.9	337	1	G3P_CURLU	O43026 curvularia
43	768.5	44.8	336	1	G3P2_ASIMG	O12552 aspergillus
44	766.5	44.7	336	1	G3P_EMENT	P20445 emerigella
45	762.5	44.5	338	1	G3P_NEUCR	P54118 neurospora

ALIGNMENTS

RESULT 1
G3P_STREY
ID G3P_STREY STANDARD; PRT; 335 AA.
AC P50467;
DT 01-OCT-1996 (Rel. 34, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH)
DR (plasma-binding protein) (plasma receptor).
GN GAP OR PLR OR GAPA OR SPY0274 OR SPY18_0261.
OS Streptococcus pyogenes, and
OC Streptococcus pyogenes (serotype M18).
OC Bacteria; Firmicutes; Lactobactiales; Streptococcaceae;
OC Streptococcus.
OK NCBI TaxID=1314, 16103;
[1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-73; 160-173 AND 186-216.
RC STRAIN=64/14;
RX MEDLINE=9235491; PubMed=1322883;
RA Lotenberg R., Broder C.C., Boyle M.D., Kain S.J., Schroeder B.L.,
RA Curtiss R. III;
RT "Cloning, sequence analysis, and expression in Escherichia coli of a
RT streptococcal plasmin receptor."
RL J. Bacteriol. 174:5204-5210(1992).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=SF370 / ATCC 700294 / Serotype M1;
RX MEDLINE=21192684; PubMed=1196296;
RA Ferretti J.J., Moshan M.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Koe B.A., McLaughlin R.;
RT "Complete genome sequence of an M1 strain of Streptococcus
RT pyogenes."
RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
[3]
RP SEQUENCE FROM N.A.
RC STRAIN=MGAS8232 / Serotype M18;
RX MEDLINE=21927593; PubMed=11917108;
RA Smoot J.C., Barbhan K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
RA Sylva G.L., Sturdevant D.E., Ricklets S.M., Porcella S.F.,
RA Pakkine L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
RA Kapur V., Daly J.A., Yeazy L.G., Musser J.M.;
RT "Genome sequence and comparative microarray analysis of serotype M18
RT group A Streptococcus strains associated with acute rheumatic fever
RT outbreaks."
RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).
[4]
RP SEQUENCE OF 1-11; 20-30; 103-128; 162-171 AND 199-215.
RC STRAIN=DRS4 / Serotype M6;
RA Hogenboom R.A., Du P., Stevenson T.I., Whitton M., Kilby G.W., Rogers J.,
RA VanBogelen R.A.;
RT "Two-dimensional gel electrophoresis map of Streptococcus pyogenes
RT proteins."
RL Submitted (MAY-2000) to Swiss-Prot.
-!- FUNCTION: ALSO BINDS HUMAN PLASMINOGEN.

DB (Plasminogen-binding protein) (plasmin receptor).
DN GAP OR PLR OR SPYM -0201 OR SPS0207.
OS Streptococcus pyogenes (serotype M3).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
CC Streptococcus.
XX NCBI_TaxID=198466;
PN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MGA5315 / Serotype M3;
RX MEDLINE=22133808; PubMed=12122206;
RA Bers S.B., SyVa G.L., Barbhan K.D., Lei B., Hoff J.S.,
RA Okahashi N., Kawabata S., Yamashita A., Nakata M., Tomiyasu Y.,
RA Hayashi H., Hattori M., Hanada S.;
RT "Genome sequence of an M3 strain of Streptococcus pyogenes reveals a
RT large-scale genomic rearrangement in invasive strains and new insights
RT into phase evolution.";
RL Genome Res. 13:1042-1055(2003).
CC -1- FUNCTION: ALSO BINDS HUMAN PLASMINOGEN (By similarity).
CC -1- CATALYTIC ACTIVITY: D-glycerinaldehyde 3-phosphate + phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -1- PATHWAY: Second phase of glycolysis; first step.
CC -1- SUBUNIT: Homotrimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
CC dehydrogenase family.

CC This SMS-S-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC CC
DR EMBL; AE014140; AAM78808.1; --
DR EMBL; AP005141; BAC63302.1; --
DR InterPro; IPR0010173; GAP_dihydrogenase.
DR InterPro; IPR006424; GAPDH-T.
DR Pfam; PF02800; gpdh_C.1.
DR Pfam; PF02800; gpdh_C.1.
DR PRINTS; PRO0078; G3PDHGRNAS.
DR TRIPFAMS; TIGR01534; GAPDH-I.1.
DR PROSITE; PS00071; GAPDH: 1.
KM Glycylsysteine oxidoreductase; NAD: Complete proteome.
FT INIT_MET 0
FT BINDING 151 151 GLYCERALDEHYDE 3-PHOSPHATE (BY
FT SIMILARITY).
FT ACT_SITE 178 178 ACTIVATES THIOL GROUP DURING CATALYSIS
FT FT (BY SIMILARITY).

SEQUENCE 335 AA; 35841 MW; 4DCB76F382FEFF698 CRC64;

Query Match 99.4%; Score 1705; DB 1; Length 335;
Best Local Similarity 99.4%; Pred. No. 6.3e-111;
Matches 333; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

2 VKVKGNGSPRIQLAFRRIQNVEGVYEVVRINDLTDPNNMLATLLKYDTQGRDGTIVEVK 61
1 VVKVGNGSPRIQLAFRRIQNVEGVYEVVRINDLTDPNNMLATLLKYDTQGRDGTIVEVK 60
62 EGGEVNGNFITKYSARDPENIDMATDGEIVLAATGTFPAKKEAAEKLALNCAKKAVTT 121

Db 61 EGGEVGNNEFKVSAERDPENIDWATDVEIVLEATGFFAKKEAEKRLHTNNAKKVIT 120
 QY 122 APGNDVKKVVFNTNHDITDGTETVYSGASCTTNCIAPARAKLHNRSGICKSMTTHAY 181
 Db 121 ARGNDVKKVVFNTNHDITDGTETVYSGASCTTNCIAPARAKLHNRSGICKSMTTHAY 180
 QY 182 TGDQMLDGEHRGSDLRARAGANIVPNSGAAKALGVPELNGKLDGAQRVPVPTG 241
 Db 181 TGDQMLDGEHRGSDLRARAGANIVPNSGAAKALGVPELNGKLDGAQRVPVPTG 240
 QY 242 SVTELVTITDKNVSVDEINAMKAASNDSPGYTEDPVSSDIVGVSGLFPAATQTKWE 301
 Db 241 SVTELVTITDKNVSVDEINAMKAASNDSPGYTEDPVSSDIVGVSGLFPAATQTKWE 300
 QY 302 VDSQQLVKVSWYDNEMSYTAQLVRLTEYFAKIAK 336
 Db 301 VDSQQLVKVSWYDNEMSYTAQLVRLTEYFAKIAK 335

RESULT 3

G3P_STREQ
 ID G3P_STREQ STANDARD; PRT; 335 AA.
 AC CS9306;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH)
 DE (Plasminogen-binding protein) (Plasmin receptor).
 GN GAP OR GAPC.
 OS Streptococcus equisimilis.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 *OC Streptococcus.
 OX NCBI_TaxId=19602;
 RX [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=H46A;
 RX MEDLINE=96305364; PubMed=8706717;
 RA Gase K., Gase A., Schirmer H., Malke H.;
 RT "Cloning, sequencing and functional overexpression of the
 RT Streptococcus equisimilis H46A gapc gene encoding a
 RT glyceraldehyde-3-phosphate dehydrogenase that also functions as a
 RT plasminogen-binding protein. Purification and biochemical
 RT characterization of the protein.";
 RL Eur. J. Biochem. 239:42-51(1996).
 CC -1- FUNCTION: ALSO BINDS HUMAN PLASMINOGEN.
 CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate +
 CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
 CC -1- PATHWAY: Second phase of glycolysis, first step.
 CC -1- SUBUNIT: Homotetramer (by similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
 CC dehydrogenase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC
 CC EMBL: X97788; CAA6377.1; -
 CC EMBL: Y12602; CAA7317.1; -
 CC PIR: S71350; S71350.
 CC HSSP: P00362; 1GDI.
 CC InterPro: IPR000173; GAP_dhrogenase.
 CC InterPro: IPR006424; GAPDH-1.
 CC Pfam: PF00044; gpdh_1.
 CC Pfam: PF02800; gpdh_C_1.
 CC PRINTS: PR00078; G3PDHGNASE.
 CC TIGRFAM: TIGR01534; GAPDH-1; 1.
 CC PROSITE: PS00071; GAPDH: 1.
 CC Glycolysis; Oxidoreductase; NAD.

FT INIT MET 0 0 BY SIMILARITY.
 FT BLINDING 151 151 GLYCERALDEHYDE 3-PHOSPHATE (BY
 FT ACT_SITE 178 178 SIMILARITY).
 FT ACT_SITE ACTIVATES THIOL GROUP DURING CATALYSIS
 SQ SEQUENCE 335 AA; 35739 MW; FE7ACFPED7663B46 CRC64;
 Query Match 94.0%; Score 1612; DB 1; Length 335;
 Best Local Similarity 95.5%; Freq. No. 1.7e-104;
 Matches 320; Conservative 2; Mismatches 13; Indels 0; Gaps 0;

QY 2 VVKVINGFGRIGRLAFRIQNVGEVYTRINDLTPPMALHLIKDTTQGRFPGTVVK 61
 Db 1 VVKVINGFGRIGRLAFRIQNVGEVYTRINDLTPPMALHLIKDTTQGRFPGTVVK 60
 QY 62 EGGEVGNNEFKVSAERDPENIDWATDVEIVLEATGFFAKKEAEKRLHTNNAKKVIT 121
 Db 61 EGGEVGNNEFKVSAERDPENIDWATDVEIVLEATGFFAKKEAEKRLHTNNAKKVIT 120
 QY 122 APGNDVKKVVFNTNHDITDGTETVYSGASCTTNCIAPARAKLHNRSGICKSMTTHAY 18-
 Db 121 APGNDVKKVVFNTNHDITDGTETVYSGASCTTNCIAPARAKLHNRSGICKSMTTHAY 180
 QY 182 TGDQMLDGEHRGSDLRARAGANIVPNSGAAKALGVPELNGKLDGAQRVPVPTG 241
 Db 181 TGDQMLDGEHRGSDLRARAGANIVPNSGAAKALGVPELNGKLDGAQRVPVPTG 240
 QY 242 SVTELVTITDKNVSVDEINAMKAASNDSPGYTEDPVSSDIVGVSGLFPAATQTKWE 301
 Db 241 SVTELVTITDKNVSVDEINAMKAASNDSPGYTEDPVSSDIVGVSGLFPAATQTKWE 300
 QY 302 VDSQQLVKVSWYDNEMSYTAQLVRLTEYFAKIAK 336
 Db 301 VDSQQLVKVSWYDNEMSYTAQLVRLTEYFAKIAK 335

RESULT 4

G3P_LACTA
 ID G3P_LACTA STANDARD; PRT; 337 AA.
 AC P52587;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH).
 GN GAP OR LD0559.
 OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
 OX NCBI_TaxId=1360;
 RX [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=LM0230;
 RX MEDLINE=95291425; PubMed=7773380;
 RA Cancelli M.R., Hillier A.D., Davidson B.E.;
 RT "Lactococcus lactis glyceraldehyde-3-phosphate dehydrogenase gene,
 RT gap: further evidence for strongly biased codon usage in glycolytic
 RT pathway genes.";
 RL Microbiology 141:1027-1036(1995).
 CC
 CC RP SEQUENCE FROM N.A.
 CC STRAIN=IL1403;
 CC MEDLINE=21235186; PubMed=11337471;
 CC Boloitin A., Wincker P., Manger S., Jallion O., Malarre K.,
 CC Weissenbach J., Ehrlich S.D., Sorokin A.;
 CC "The complete genome sequence of the lactic acid bacterium Lactococcus
 CC lactis ssp. lactis IL1403.";
 CC Genome Res. 11:731-753(2001).
 CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
 CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
 CC -1- PATHWAY: Second phase of glycolysis, first step.
 CC -1- SUBUNIT: Homotetramer (by similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
 CC dehydrogenase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; 136307; AAC1453.1; -
DR EMBL; AEO06290; AAK04657.1; -
DR PIR; G86694; G86694.
DR HSSP; P17721; 1HDG.
DR InterPro; IPR00173; GAP dhdtrogenase.
DR InterPro; IPR006424; GAPDH-1.
DR Pfam; PF00044; spdh.1.
DR Pfam; PF02800; spdh.C.1.
DR PRINTS; PR00078; G3PDHGNASE.
DR TIGRFAMs; TIGR01534; GAPDH-1; 1.
DR PROSITE; PS00071; GAPDH; FALSE NEG.
KW Glycolysis; Oxidoreductase; NAD; Complete proteome.
FT BINDING 152 152 GLYCERALDEHYDE 3-PHOSPHATE
FT ACT SITE 179 179 (BY SIMILARITY).
FT CONFLICT 143 143 ACTIVATES THIOL GROUP DURING CATALYSIS
FT (BY SIMILARITY).
SQ SEQUENCE 337 AA; 36057 MW; 17BB8C6A8FEF589D CRC64;

Query Match 78.0%; Score 1338.5; DB 1; Length 337;
Best Local Similarity 77.2%; Pred. No. 1.4e-85;
Matches 260; Conservative 31; Mismatches 45; Indels 1; Gaps 1;

QY 1 MTKVKGINGEGRIGRLAFRIQNEVEYTRINDLTDPMALHLKYDTQGRPGTVEV 60
DB 1 MTKVKGINGEGRIGRLAFRIQNEVEYTRINDLTDPMALHLKYDTQGRPGTVEV 60
QY 61 KEGEYVNGFIVKASERDENIDMADVEIVLAEATGFAKKEAAEKHLHANGAKKVTI 120
DB 61 KEDGPDVNGKVFYVFAERPEIDQADSGVEIVLAEATGFAKKEAAEKHLHANGAKKVTI 120
QY 121 TAPGNDVKTWVNTNHDILDTETVTSAGSCTTNCCLAPMAKALHDAFQIGLMTTHA 180
DB 121 TAPGNDVKTWVNTNHDILDTETVTSAGSCTTNCCLAPMAKALHDAFQIGLMTTHA 180
QY 181 YTGDMTLDGPHRGDLRRARAGANIVPNSGAKAIGLVIPELNGKLDGAAGVPT 240
DB 181 YTGDMTLDGPHRGDLRRARAGANIVPNSGAKAIGLVIPELNGKLDGAAGVPT 240
QY 241 GSVEELVYTLDDKNSVDEINAMKAASNDSPGYTDPVSSDIVGYSGLFDATQTKM 300
DB 241 GSVEELVYTLDDKNSVDEINAMKAASNDSPGYTDPVSSDIVGYSGLFDATQTKM 300
QY 301 EY-DGSQLVKTWVNTNHDILDTETVTSAGSCTTNCCLAPMAKALHDAFQIGLMTTHA 360
DB 301 EY-DGSQLVKTWVNTNHDILDTETVTSAGSCTTNCCLAPMAKALHDAFQIGLMTTHA 360
QY 301 DKGQGLVKTWVNTNHDILDTETVTSAGSCTTNCCLAPMAKALHDAFQIGLMTTHA 360
DB 301 DKGQGLVKTWVNTNHDILDTETVTSAGSCTTNCCLAPMAKALHDAFQIGLMTTHA 360

RESULT 5
G3P_CLOAB STANDARD; PRT; 334 AA.
ID G3P_CLOAB
AC Q59309;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH) (CP
DE 17/CP 18).
GN GAP.
OS Clostridium pasteurianum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1501;
RX [1]
RP SEQUENCE FROM N.A.

RA Oster T., Assochei O., Scherrer S., Branlant G., Branlant C.;
RL Submitted (JUN-1993) to the EMBL/GenBank/DBJ databases.
RN [2]
RP PARTIAL SEQUENCE OF 1-26.
RC STRAIN=W5;
RX MEDLINE=98291870; PubMed=9629918;
RA Flensburg R., Skjeldal L.;
RT "Two-dimensional gel electrophoresis separation and N-terminal
RT sequence analysis of proteins from Clostridium pasteurianum W5";
RL Electrophoresis 15:802-806(1998).
CC -1 CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -1 PATHWAY: Second phase of glycolysis; first step.
CC -1 SUBUNIT: Homotrimer (by similarity).
CC -1 SUBCELLULAR LOCATION: Cytoplasmic.
CC -1 SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
CC dehydrogenase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; X72219; CA451020.1; -
DR PIR; S34254; S34254.
DR HSSP; P00362; 1GDI.
DR InterPro; IPR00173; GAP dhdtrogenase.
DR InterPro; IPR006424; GAPDH-1.
DR Pfam; PF00044; spdh.1.
DR Pfam; PF02800; spdh.C.1.
DR PRINTS; PR00078; G3PDHGNASE.
DR TIGRFAMs; TIGR01534; GAPDH-1; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW Glycolysis; Oxidoreductase; NAD.
FT BINDING 150 150 GLYCERALDEHYDE 3-PHOSPHATE.
FT ACT SITE 177 177 ACTIVATES THIOL GROUP DURING CATALYSIS.
SQ SEQUENCE 334 AA; 36078 MW; D1590SD0DA76227 CRC64;

Query Match 75.2%; Score 1289.5; DB 1; Length 334;
Best Local Similarity 74.3%; Pred. No. 3.4e-82;
Matches 249; Conservative 33; Mismatches 52; Indels 1; Gaps 1;

QY 2 VTKVKGINGEGRIGRLAFRIQNEVEYTRINDLTDPMALHLKYDTQGRPGTVEV 61
DB 1 MTKVKGINGEGRIGRLAFRIQNEVEYTRINDLTDPMALHLKYDTQGRPGTVEV 60
QY 61 KEGEYVNGFIVKASERDENIDMADVEIVLAEATGFAKKEAAEKHLHANGAKKVTI 121
DB 61 EGAFFVNGKVFYVFAERPEIDQADSGVEIVLAEATGFAKKEAAEKHLHANGAKKVTI 119
QY 122 APGANDVKTWVNTNHDILDTETVTSAGSCTTNCCLAPMAKALHDAFQIGLMTTHA 181
DB 122 APGANDVKTWVNTNHDILDTETVTSAGSCTTNCCLAPMAKALHDAFQIGLMTTHA 179
QY 180 TNDQMTLDGPHRGDLRRARAGANIVPNSGAKAIGLVIPELNGKLDGAAGVPT 241
DB 180 TNDQMTLDGPHRGDLRRARAGANIVPNSGAKAIGLVIPELNGKLDGAAGVPT 239
QY 242 SVTEELVYTLDDKNSVDEINAMKAASNDSPGYTDPVSSDIVGYSGLFDATQTKM 301
DB 240 SVTEELVYTLDDKNSVDEINAMKAASNDSPGYTDPVSSDIVGYSGLFDATQTKM 299
QY 302 VDSGQLVKTWVNTNHDILDTETVTSAGSCTTNCCLAPMAKALHDAFQIGLMTTHA 366
DB 300 VDSGQLVKTWVNTNHDILDTETVTSAGSCTTNCCLAPMAKALHDAFQIGLMTTHA 334

RESULT 6
G3P_CLOAB STANDARD; PRT; 334 AA.
ID G3P_CLOAB
AC Q59309;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH) (CP
DE 17/CP 18).
GN GAP.
OS Clostridium pasteurianum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1501;
RX [1]
RP SEQUENCE FROM N.A.

AC 052631;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH).
 GN GAP OR GAPC OR GAC0709.
 OS Clostridium acetobutylicum.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxId=1488;
 RN [1]
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
 RX MEDLINE=99392446; PubMed=10463150;
 RA Schreiber W., Durie P.;
 RA "The glyceraldehyde-3-phosphate dehydrogenase of Clostridium
 RT acetobutylicum: isolation and purification of the enzyme, and
 RT sequencing and localization of the gap gene within a cluster of other
 RT glycolytic genes.";
 RL Microbiology 145:1839-1847(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
 RX MEDLINE=21359325; PubMed=11466286;
 RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
 RA Gibson R., Lee H.W., Dubois J., Qiu D., Hittl J., Wolf Y.I.,
 RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.T.,
 RA Bennett G.N., Koonin E.V., Smith D.R.;
 RA "Genome sequence and comparative analysis of the solvent-producing
 RT bacterium Clostridium acetobutylicum.";
 RL J. Bacteriol. 183:4823-4838(2001).
 CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
 CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
 CC -1- PATHWAY: Second phase of glycolysis; first step.
 CC -1- SUBUNIT: Homotrimer.
 CC -1- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
 CC dehydrogenase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF043386; AAC13160.1; -;
 DR EMBL: AE007586; AAK78686.1; -;
 DR PIR: C36987; C36987.
 DR HSSP: P17721; 1HDG.
 DR InterPro: IPR000173; GAP_dhdrogenase.
 DR InterPro: IPR006424; GAPDH-I.
 DR Pfam: PF00044; Spdh. 1.
 DR Pfam: PF02800; Spdh. C. 1.
 DR PRINTS: PR00078; G3PDHDSGNASE.
 DR TIGRFAMs: TIGR01534; GAPDH-I. 1.
 DR PROSITE: PS00071; GAPDH. 1.
 DR GlycoLysis: Oxidoreductase; NAD: Complete proteome.
 FT ACT_SITE 150 150 GLYCERALDEHYDE 3-PHOSPHATE.
 FT BINDING 177 177 ACTIVATES THIOL GROUP DURING CATALYSIS.
 SQ SEQUENCE 334 AA; 35850 MW; 10C52A174BE789B5 CRC64;
 Query Match 73.4%; Score 1259.5; DB 1; Length 334;
 Best Local Similarity 71.9%; Pred. No. 4e-80;
 Matches 241; Conservative 40; Mismatches 53; Indels 1; Gaps 1;
 QY 2 VYKVGINGFGRIGRLAFRIQVNGVEYTRINDLTPNNALHLKYDTGGRFGDGVYK 61
 DB 1 MAKIINGFGIRIALRILREVPGLLEVAININDLTKAKIALHLFKYDGSQGRFNGSIRV 60
 QY 62 EGGEVNGNPIKVSARDEPNIDWATGCGVIVELANGPRAKKAEEHILAMAKKVVIT 121

DB 61 BGAFVNGKEVFAEADPEKLPNGDLIDVLECTGFEFFKKEKAEANVA-GAKKVIIS 119
 QY 122 AGGQDKYTPVNTVNDHLLDGETVIVSGSCCTNCLAPMAKALDARGLOKIMTTIHAY 181
 DB 120 APAGSDKTIIVNNWNEEDGETVIVSGSCCTNCLAPMAKALDARGLOKIMTTIHAF 179
 QY 182 TGDQMILDGPHRQZDLRRARAGANIVPNSGAAVAIGLVIPIRNGKLDGAAQRPVPTG 241
 DB 180 TNDQNTLDGPHRQZDLRRARAGANIVPNSGAAVAIGLVIPIRNGKLDGAAQRPVPTG 239
 QY 242 SVTELVTTLDKRVSDVEINAAKASNDSPFTEDPIYSDIVGVSGSLFPAITQXME 301
 DB 240 SITELVSLKKKVVYEEINAAKASNDSPFTEDPIYSDIVGVSGSLFPAITQXME 299
 QY 302 VDSGLKVVSWTDNBSNTYQVLETLVEFATIAK 336
 DB 300 VNSQLVTKAAMDWMSYTSQVRLTALYFAKIAK 334
 RESULT 7
 G3PL STRAP STANDARD; PRT; 336 AA.
 ID G3PL STRAP STANDARD; PRT; 336 AA.
 AC Q8CPY5;
 DT 15-MAR-2004 (Rel. 43, Created)
 DT 15-MAR-2004 (Rel. 43, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Glyceraldehyde 3-phosphate dehydrogenase 1 (EC 1.2.1.12) (GAPDH 1).
 GN GAP OR GAP OR SE0557.
 OS Staphylococcus epidermidis.
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 OX NCBI_TaxId=1282;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 12228;
 RX PubMed=12850922;
 RA Zhang Y.-Q., Ren S.-X., Li H.-L., Wang Y.-X., Fu G., Yang J.,
 RA Qin Z.-Q., Mao Y.-G., Wang W.-Y., Chen R.-S., Shen Y., Chen Z.,
 RA Yuan Z.-H., Zhao G.-P., Qu D., Danchin A., Wen Y.-M.;
 RA "Genome-based analysis of virulence genes in a non-biofilm-forming
 RT Staphylococcus epidermidis strain (ATCC 12228).";
 RL Mol. Microbiol. 49:1577-1593(2003).
 CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
 CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
 CC -1- PATHWAY: Second phase of glycolysis; first step.
 CC -1- SUBUNIT: Homotrimer (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
 CC dehydrogenase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AE016745; AA004154.1; -;
 DR InterPro: IPR006424; GAP_dhdrogenase.
 DR InterPro: IPR000173; GAP_dhdrogenase.
 DR Pfam: PF00044; Spdh. 1.
 DR Pfam: PF02800; Spdh. C. 1.
 DR PRINTS: PR00078; G3PDHDSGNASE.
 DR TIGRFAMs: TIGR01534; GAPDH-I. 1.
 DR PROSITE: PS00071; GAPDH. 1.
 DR GlycoLysis: Oxidoreductase; NAD: Complete proteome.
 FT BINDING 151 151 GLYCERALDEHYDE 3-PHOSPHATE.
 FT ACT_SITE 178 178 SIMILARITY.
 FT ACTIVATES THIOL GROUP DURING CATALYSIS.
 FT (BY SIMILARITY).
 SQ SEQUENCE 336 AA; 36190 MW; A96220202A5767 CRC64;
 Query Match 69.4%; Score 1190; DB 1; Length 336;

DT 28-FEB-2003 (Rel. 41, last annotation update)
 DE Glyceraldehyde 3-phosphate dehydrogenase [EC 1.2.1.12] (GAPDH).
 GN GAP.
 OS Lactobacillus delbrueckii (subsp. bulgaricus).
 OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
 CC Lactobacillus.
 CX NCBI_TaxId=1585;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BI07;
 RX MEDLINE=98240227; PubMed=9579064;
 RA Branny P., Delatorre F., Garel J.R.;
 RT "An operon encoding three glycolytic enzymes in Lactobacillus
 RT delbrueckii subsp. bulgaricus: glyceraldehyde-3-phosphate
 RT dehydrogenase, phosphoglycerate kinase and triosephosphate
 RT isomerase." J. 144:905-914 (1998).
 RL Microbiology. 144:905-914 (1998).
 CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
 CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
 CC -1- PATHWAY: Second phase of glycolysis; first step.
 CC -1- SUBUNIT: Homotetramer (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
 CC dehydrogenase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; AJ000339; CAA04014.1; -
 DR PIR; T09633; T09633.
 DR HSSP; P17721; IHG.
 DR InterPro; IPR000173; GAP_dhndrogenase.
 DR InterPro; IPR006424; GAPDH-I.
 DR Pfam; PF00044; gpdh_C.1.
 DR Pfam; PF02800; gpdh_C.1.
 DR PRINTS; PR00078; G3PDHGRNAS.
 DR TIGRFAMs; TIGR01534; GAPDH-I.1.
 DR PROSITE; PS00071; GAPDH; FALSE NEG.
 DR KEGG; glycolysis; NAD; Oxidoreductase.
 FT BINDING 156 156 GLYCERALDEHYDE 3-PHOSPHATE
 FT (BY SIMILARITY).
 FT ACT_SITE 183 183 ACTIVATES THIOL GROUP DURING CATALYSIS
 FT (BY SIMILARITY).
 FT SEQUENCE 338 AA; 36564 MW; DF978C9CA4F7DFA CRC64;
 Query Match 58.5%; Score 1003; DB 1; Length 338;
 Best Local Similarity 58.7%; Pred. No. 2,26-62;
 Matches 193; Conservative 45; Mismatches 89; Indels 6; Gaps 3;
 1 MVAKVG:INGFGRIGLAFRI---ONVEGEVTRINDLTPDMLHLKTYDTQGFDS 56
 1 MTKGIGNGFGRIGLAFRIIMJGSETDIEVAINDLTPPMLHLKTYDTQGFDS 60
 57 TVEVKGCFEYNGNFIKYSERDPEITMA-TDGEYLYEATGFPAKKAEEKILANNA 115
 61 EVSATSDSLVVGKKYRYVAEPQANIPWKKNDGDEVLECTGFYSYAKAGQNTIDA-GA 119
 116 KKVITAPGNDYKVVVNNHDLIGETVYSGASCTTGCLAPAKALHAPGIGQGLM 175
 120 KKVILAPAGNDLKTIVSYVNOTLTADDTIYSGASCTTSLAPANALKEFGIYQGTW 179
 176 TTTHAYTGDMITLDPHNGDLRPARAGANIVPNSGAAKALGIVTELGKIDGAQR 235
 180 TTTHAYTATQKVLDPGRKNNPNNAAALENIIPHSYGAKAIGIVTELGKIDGAQR 239
 236 VVPVTSVTEIVITLDKQVSDVDEINAAKASNSFGITDPIYSVDIVGVSQSLFDAT 295
 240 VPVDSSETELVITLDEKKVTAEEVNAAMKRYESBSPAFADQIVTDVIGTASIDPI 299

QY 296 OTKWEKDSQLYKVVSWYDNEWSTYATQVLTLEYPAKI 334
 DB 300 QTVITAEKQLYKVVYAWYDNEWSTYATQVLTLEYPAKI 338
 RESULT 10
 ID GAP3_ECO57 STANDARD; PRT; 333 AA.
 AC P58072;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Glyceraldehyde 3-phosphate dehydrogenase C [EC 1.2.1.12] (GAPDH-C).
 GN GAPC OR 22304 OR BCS2022.
 OS Escherichia coli O157:H7.
 CC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 CC Enterobacteriaceae; Escherichia.
 CX NCBI_TaxId=83334;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
 RX MEDLINE=21074935; PubMed=11206551;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor V., Kirkpatrick N.A.,
 RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamoudis K.,
 RA Apodaca J., Nantharajan T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RT "Genome sequence of enterohemorrhagic Escherichia coli O157:H7.";
 RL Nature 409:529-533 (2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / RIMD 050952;
 RX MEDLINE=21156231; PubMed=11258796.
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
 RA Kihara S., Shiba T., Hattori M., Shinagawa H.;
 RT "Complete genome sequence of enterohemorrhagic Escherichia coli
 RT O157:H7 and genomic comparison with a laboratory strain K-12.";
 RL DNA Res. 8:111-22 (2001).
 CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
 CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
 CC -1- PATHWAY: Second phase of glycolysis; first step.
 CC -1- SUBUNIT: Homotetramer (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
 CC -1- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
 CC dehydrogenase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; AB005364; AA056359.1; -
 DR EMBL; AF002557; BAB35445.1; -
 DR PIR; C85737; C85737.
 DR PIR; F90881; F90881.
 DR HSSP; P17721; IHG.
 DR InterPro; IPR000173; GAP_dhndrogenase.
 DR InterPro; IPR006424; GAPDH-I.
 DR Pfam; PF00044; gpdh_C.1.
 DR Pfam; PF02800; gpdh_C.1.
 DR PRINTS; PR00078; G3PDHGRNAS.
 DR TIGRFAMs; TIGR01534; GAPDH-I.1.
 DR PROSITE; PS00071; GAPDH; 1.
 DR KEGG; glycolysis; NAD; Multigene family; Complete proteome.
 FT BINDING 150 150 GLYCERALDEHYDE 3-PHOSPHATE
 FT (BY SIMILARITY).

DR Pfam: PF02800; gpdh C1.
 DR PRINTS: PR00078; G3PDHDEGNASE.
 DR TIGRPFAM: TIGR01534; GAPDH-1; 1.
 DR PROSITE: PS00071; GAPDH; 1.
 DR KEGG: Glycolysis; Oxidoreductase; NAD; Multigene family; Complete proteome.
 FT BINDING 150 150
 FT ACT_SITE 177 177
 FT CONFICT 39 39
 FT SEQUENCE 333 AA; 35649 MW; E88223297376BDA0 CRC64;
 Query Match 57.8%; Score 992; DB 1; Length 333;
 Best Local Similarity 59.3%; Pred. No. 12e-61;
 Matches 199; Conservative 54; Mismatches 76; Indels 6; Gaps 5;

QY 4 KVGINGFGRIGRLAERRIQNVE-GVEVTRINDLTPNNLAHLKYDTGGRPGTVEYKE 62
 DB 3 KVGINGFGRIGRLVGRLEVKSNIDVVALINDLSKRIIAYLLKSDNSYGPFPMSYDTE 62
 QY 63 GGEFVNGFIRVSAERDPENIDMTDGVETVLTATGTFPAKKAERKHLANAKKVVITA 122
 DB 63 DSLIYDGKSLVYAEKAKNIPWAKAGAEIIVCTGPTYSAKESQAHIDA-GAKKVLISA 121
 QY 123 PGANDVKTIVFNTNHDLDGTEVTISGASCTTNCIALPAAALDAGLGKXGMMTHIAYT 182
 DB 122 PAG-EMKTIYVNVDDTLGDNITIVSASCTTNCIALPAAALHDSFGIEVGTTHIAYT 180
 QY 183 GQDMILDPHRRGDLRRARAGANIVPSTGAKAIGLVIPELNGKLDGAAQVPPVPTGS 242
 DB 181 GQSLVDPG-RGKDLRASEPAAENIIPHTGAAGAIGLVIPELNGKLDGAAQVPPVPTGS 239
 QY 243 VTEIYVTLDRKVSDEINAMAKKA-SNDSGCTEDPTYSSTIVSGSLPATOTKYM 300
 DB 240 VTEIYVTLGKVKYAEVNNALRKQATNNESGTYDEIVSSDISSHSGVFPATQET 299
 QY 301 EYDGSQVYVSWYDNEMSYTAQLVETLEYPAKI 334
 DB 300 AVGDQLYKVTAMWINDNYSYVQLIRTEKPAKI 333

RESULT 12
 G3P_MYCPN STANDARD; PRT; 337 AA.
 ID G3P_MYCPN
 AC PF0559;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH).
 GN GAP OR GAP OR MPN430 OR MP411.
 OS Mycoplasma pneumoniae.
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID=2104;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 29342 / M129;
 RX MEDLINE=91105885; PubMed=8948633;
 RA Himmelreich R., Hilbert H., Plaegens H., Finkl E., La B.-C.,
 RA Herrmann R.;
 *Complete sequence analysis of the genome of the bacterium Mycoplasma
 pneumoniae.";
 RT Nucleic Acids Res. 24:4420-4449 (1996).
 CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
 NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
 CC -1- PATHWAY: Second phase of glycolysis; first step.
 CC -1- SUBUNIT: Homotetramer (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: belongs to the glyceraldehyde 3-phosphate
 dehydrogenase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AE000040; AAB96059.1; -
 DR PIR: S73737; S73737.
 DR HSSP: P17721; IHOG.
 DR InterPro: IPR000173; GAP_dhrogenase.
 DR InterPro: IPR006424; GAPDH-1.
 DR Pfam: PF00044; gpdh C1.
 DR Pfam: PF02800; gpdh C1.
 DR PRINTS: PR00078; G3PDHDEGNASE.
 DR TIGRPFAM: TIGR01534; GAPDH-1; 1.
 DR PROSITE: PS00071; GAPDH; 1.
 DR Glycolysis; Oxidoreductase; NAD; Complete proteome.
 FT BINDING 157 157
 FT ACT_SITE 184 184
 FT SEQUENCE 337 AA; 36805 MW; 550747A529ABCAB3 CRC64;
 Query Match 53.7%; Score 921; DB 1; Length 337;
 Best Local Similarity 54.1%; Pred. No. 1e-56;
 Matches 180; Conservative 54; Mismatches 95; Indels 4; Gaps 3;

QY 3 KVGINGFGRIGRLAERRIQNVEGVEVTRINDLTPNNLAHLKYDTGGRPGTVEYKE 62
 DB 8 IRVAINRFGIRGLVFRALSSQKRIEIVANVNDLHPDLAHLKYSANGCEFFKKVAVKD 67
 QY 63 GGEFVNGFIRVSAERDPENIDMTDGVETVLTATGTFPAKKAERKHLANAKKVVITA 122
 DB 68 KTLIMIDKKVIVSESDANLPAAENIIVESTGRFVSEEGASLHLDA-GAKKVIISA 126
 QY 123 PG-ANDVKTIVFNTNHDLDGTEVTISGASCTTNCIALPAAALDAGLGKXGMMTHIAY 181
 DB 127 PAKKTIYVNVVWNTKTNMEDTKISASCTTNCIALPAAALVLEKPNGLHGMTHIAY 186
 QY 182 TGDQMLDPHRRGDLRRARAGANIVPSTGAKAIGLVIPELNGKLDGAAQVPPVPTG 241
 DB 187 TRDQRLDPAH--SDRRARRAANIVPTTGAAGAIGLVIPELNGKLDGAAQVPPVPTG 244
 QY 242 SVTEIYVTLDRKVSDEINAMAKKA-SNDSGCTEDPTYSSTIVSGSLPATOTKYM 301
 DB 245 SIYELCVALKEDKAWEOINQAKASAPRCEDEIVSSDISSHSGVFPATQET 304
 QY 302 VDGSLYVYVSWYDNEMSYTAQLVETLEYPAKI 334
 DB 305 VDGKRLYKVTAMWINDNYSYVQLIRVAVYCAKI 337

RESULT 13
 G3P_CORGL STANDARD; PRT; 334 AA.
 ID G3P_CORGL
 AC Q01651;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH).
 GN GAP OR CG1588.
 OS Corynebacterium glutamicum (Brevibacterium flavum).
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
 OX NCBI_TaxID=1718;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 13059 / AS019;
 RX MEDLINE=93015645; PubMed=1400158;
 RA Birkmanns B.J.;
 *Identification, sequence analysis, and expression of a
 RT Corynebacterium glutamicum gene cluster encoding the three glycolytic
 RT enzymes glyceraldehyde-3-phosphate dehydrogenase, 3-phosphoglycerate
 RT kinase, and triosephosphate isomerase.";

RL J. Bacteriol. 174:6076-6086(1992).
 RN [2]
 RC SEQUENCE FROM N.A.
 RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
 RA Nakagawa S.;
 RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
 RT Submitted (MAY-2002) to the EMBL/Genbank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
 CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
 CC -1- PATHWAY: Second phase of glycolysis; first step.
 CC -1- SUBUNIT: Homotetramer.
 CC -1- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
 CC dehydrogenase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

CC EMBL; X59403; CAA42045.1; -;
 DR EMBL; AP005279; BAB98981.1; -;
 DR PIR; A43260; A43260.
 DR HSSP; P00362; 1GDI.
 DR InterPro; IPR00173; GAP dehydrogenase.
 DR InterPro; IPR006424; GAPDH-1.
 DR Pfam; PF00344; spdh; 1.
 DR Pfam; PF02800; spdh C; 1.
 DR PRINTS; PR00078; G3EDHGNASE.
 DR TIGRFSMS; TIGR01534; GAPDH-1; 1.
 DR PROSITE; PS00071; GAPDH; 1.
 CC GLYCOLYSIS; Oxidoreductase; NAD; Complete proteome.
 CC BINDING 153 153 GLYCERALDEHYDE 3-PHOSPHATE.
 FT ACT SITE 180 180 ACTIVATES THIOL GROUP DURING CATALYSIS.
 FT CONFLICT 25 26 SD -> NG (IN REF. 1).
 FT CONFLICT 333 334 KL -> QALN (IN REF. 1).
 SQ SEQUENCE 334 AA; 36045 MW; 33792AF65FA90F7 CRC64;

Query Match 53.0%; Score 908.5; DB 1; Length 334;
 Best Local Similarity 56.2%; Pred. No. 7.4e-56;
 Matches 152; Conservative 41; Mismatches 94; Indels 13; Gaps 6;

QY 1 MMYKGVNGPGRIGLAERLI QNVEGVYRINDLIDPMALHLKYDTQGRFDGIVE 59
 DB 1 MTRVINGPFRIGRNFPRVAVERSDDLEVAANDLTKTSTLLNFDISIMGLQZVE 60
 QY 60 VKEGFEVNGNFIKVSAREPENIDATDGEIVLEATGFEAKKEAKELHANGAKKV 119
 DB 61 YDDSDITVGGRIAYVAREPDKLMAAHNDIVIESTGFTDAAVAHIEA-GAKKV 119
 QY 120 ITAPGNDVYVFNTHD-LD-GTEIVISAGSTTNCIAEPAKALHADRISQKMTTI 178
 DB 120 ISAPASNEDEATFVGVNHSYDEHNHNTISGASCTTNCIAEPAKALHADRISQKMTTI 179
 QY 179 HATGDMITLGGPRGGDLARRAAGANIVNSGAAXIAGVTEELNGKIDGAARVY 238
 DB 180 HATGDMITLGGPRGGDLARRAAGANIVNSGAAXIAGVTEELNGKIDGAARVY 237
 QY 239 PEGSVTELVITLDKNVSVDEINAKMAASNDSPG---YTEDPIVSSDIVGVSYGIFDA 294
 DB 238 ITGSAITDLFNTKSEVYVESINAKIKBAVGEFETLAVSEBELVSTDIVHDSGISIFDA 297
 QY 295 TQTKVNEVDGSQLVVSQYDNEMSTYLAQVITILEYFA 332
 DB 298 GLTKV---SGNTVYVVSQYDNEMSTYLAQVITILEYFA 331

RESULT 14
 GAP_MCOGE STANDARD; PCT; 337 AA.
 ID GAP_MCOGE

AC P47543;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH).
 GN GAPD OR GAP OR MG301.
 OS Mycoplasma genitalium.
 CX Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
 NCBI_TaxID=2097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 33530 / G-37;
 RX MEDLINE=96026346; PubMed=7569993;
 RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
 RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
 RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhmann J.L.,
 RA Nguyen D.T., Uetshack T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
 RA Tomb J.-F., Dougherty B.A., Bock K.F., Hu P.-C., Luetter T.S.,
 RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
 RT "The minimal genome complement of Mycoplasma genitalium.";
 RL Science 270:1397-403(1995).
 RN [2]
 RP SEQUENCE OF 1-81 AND 279-337 FROM N.A.
 RC STRAIN=ATCC 33530 / G-37;
 RX MEDLINE=94075230; PubMed=8253680;
 RA Peterson S.N., Hu P.-C., Bock K.F., Hutchison C.A. III;
 RT "A survey of the Mycoplasma genitalium genome by using random
 RT sequencing.";
 RL J. Bacteriol. 175:7918-7930(1993).
 CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
 CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
 CC -1- PATHWAY: Second phase of glycolysis; first step.
 CC -1- SUBUNIT: Homotetramer (By similarity).
 CC -1- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
 CC dehydrogenase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

QY 3 VVGVNGPGRIGLAERLI QNVEGVYRINDLIDPMALHLKYDTQGRFDGIVEYKE 62
 DB 8 KVVAINGPGRIGLVERSLSKANVAVVAILNDLTPFVLAHLTKYDSANGELKRIYTKVQ 67
 Query Match 52.9%; Score 908; DB 1; Length 337;
 Best Local Similarity 53.2%; Pred. No. 8.2e-56;
 Matches 177; Conservative 54; Mismatches 98; Indels 4; Gaps 3;

QY 63 GGEVNFNFIKVAERDPNDIMATDVEIVLATGTFPAKKAEMKHLHANGAKKVVITA 122
 DB 68 NIIQIDREKVVVFSEGDPONLPMDBHDLDIVISITRPFVSEGSIMHKA-GAKKVTISA 126
 QY 123 PG-GNDVATVFENTNDIJDGETVYISGASCTTNCIAPAKALHDAFGIOKIMTTIAY 181
 DB 127 PAKEKIRIVVYNNVHTTISDDKIISAACTTNCIAPVHLLEKNEGIYVGTMLTVHAY 186
 QY 182 TSDQMIIDGPHRGDULRRARAGANIVNNGAAXIGVITPELNGKIDGAAQRPVPTG 241
 DB 187 TMDQRIQDAEPH-NDLRARAAAANIVITTTGAAKAIGLVEANGLKNGSLKRVPLTG 244
 QY 242 SATELVTELDGNVSVDEINAAKMAASNDSPGTEDEPIVSSDVGVSGLFDATQTKWE 301
 DB 245 STVELSVLEHSVPSVEVQNAKRRASAPFYCEDPIVSDVSVSEIGSIFDSKLTNIYE 304
 QY 302 VDSGQIVXVSWYNDENSYTQALVTELEFYAKI 334
 DB 305 VDSMKLYVYVAMYNESVYHQLVTVVSYCAKL 337

RESULT 15
 G3P1_BACSU STANDARD: PRT; 334 AA.
 AC P09124;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Glyceraldehyde 3-phosphate dehydrogenase 1 (EC 1.2.1.12) (GAPDH) (NMD-
 dependent glyceraldehyde 3-phosphate dehydrogenase).
 GN GABA OR GAP OR HS133940.
 CS Bacillus subtilis.
 CC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 CX NCBI Tax-ID=1423;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168 / BD170;
 RX MEDLINE=89160255; PubMed=2493623;
 RA Viene A., Dhase P.;
 RT "Sequence of the glyceraldehyde-3-phosphate dehydrogenase gene from
 Bacillus subtilis";
 RL Nucleic Acids Res. 17:1251-1251(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=86044033; PubMed=3384377;
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Berteiro M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Boriss R., Bourcier L., Brans A., Braum M., Brignell S.C., Bron S.,
 RA Brouillet S., Bruchet C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Comerton I.F., Cummings N.J., Daniel R.A.,
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Enlian K.D., Strington J., Fabre C., Ferrari E., Foulger D.,
 RA Fritz C.Y., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Gilm S.Y., Glaeser P., Goffeau A., Goldlighty E.J., Grandi G.,
 RA Giuseppe G., Guy B.J., Haga K., Haeuch J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holstapel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Lader U., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Mamel C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
 RA Moore D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
 RA Preecan B., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Rocha E., Kocbe R., Rose W., Sadate I.,
 RA Sato T., Scanlan E., Schleich S., Schroeder R., Scofield F.,
 RA Seikuchi J., Sekowska A., Serot S.J., Serro P., Shin B.S., Sollo B.,
 RA Sorokin A., Taccini B., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tamakoshi A., Tanaka T., Tempstra P., Tononelli A.,
 RA Tosi V., Uchiyama S., Vandendol M., Vannier F., Vasseroiti A.,
 RA Visi A., Wambutt R., Wedler E., Wedler H., Weitznagel T.,
 RA Winters P., Wipat A., Yamamoto A., Yamane K., Yasunoto K., Yata K.,

RA Yoshida K., Yoshikawa H.F., Zimstein E., Yoshikawa H., Darchin A.;
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 subtilis";
 RL Nature 390:249-256(1997).
 RN [3]
 RP SEQUENCE OF 1-30.
 RC STRAIN=168 / JH642;
 RX MEDLINE=96345629; PubMed=8755892;
 RA Graumann P., Schroeder K., Schmid R., Marahel M.A.;
 RT "Cold shock stress-induced proteins in Bacillus subtilis";
 RL J. Bacteriol. 178:4611-4619(1996).
 RN [4]
 RP CHARACTERIZATION;
 RX MEDLINE=20261518; PubMed=10799476;
 RA Fillingner S., Bockel-Muller S., Azza S., Deryn E., Brantant G.,
 RA Americh S.;
 RT "Two glyceraldehyde 3-phosphate dehydrogenases with opposite
 physiological roles in a nonphotosynthetic bacterium";
 RL J. Biol. Chem. 275:14031-14037(2000).
 CC -1- FUNCTION: More active in catabolism.
 CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
 NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
 CC -1- PATHWAY: Second phase of glycolysis; first step.
 CC -1- SUBUNIT: Homotetramer.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
 dehydrogenase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.ebi.ac.uk/announce/>
 or send an email to license@ebi-sib.ch).
 CC -----
 CC EMBL: X13011; CAJ31434.1; -
 DR EMBL: Z99121; CAJ15399.1; -
 DR PIR: S02754; DEBSG.
 DR HSP: P00362; IGD1.
 DR Subtilist; BGI0827; g3pA
 DR Interpro; IPR000173; GAP_dhrogenase.
 DR Interpro; IPR006424; GAPDH-1.
 DR Pfam; PF00044; gpdh; 1.
 DR Pfam; PF02800; gpdh; 1.
 DR PRINTS; PR00078; G3PDHGNASE.
 DR TIGRFBAS; TIGR01534; GAPDH-1; 1.
 DR PROSITE; PS00071; GAPDH; 1.
 KM Glycolysis; Oxidoreductase; NAD; Multigene family; Complete proteome.
 FT INIT MET 0
 FT BINDING 151 151 GLYCERALDEHYDE 3-PHOSPHATE.
 FT ACT SITE 178 178 ACTIVATES THIOLO GROUP DURING CATALYSIS.
 SQ SEQUENCE 334 AA; 35701 MW; 1283D3B6CF5095EC CRC64;

Query Match 52.0%; Score 891.5; DB 1; Length 334;
 Best Local Similarity 54.5%; Pred. No. 1.1e-54;
 Matches 103; Conservative 41; Mismatches 101; Indels 11; Gaps 5;

QY 3 YVVGINGFGHIGLAFRIQWEGVEYTRINDITPNNIMAHLLKYDTTQGRFGDGYEYKE 62
 DB 2 VVVGINGFGHIGNVRALNNPEVEVAVNNDITANNLHLQYDVSQKDAEVSVDG 61
 QY 63 GGEVNFNFIKVAERDPNDIMATDVEIVLATGTFPAKKAEMKHLHANGAKKVVITA 122
 DB 62 NNLVYNGKTEIVSAERDPKLSMKQGEIVLVSSTGFTKADAKKLEA-GAKKVTISA 120
 QY 123 PGNDVATVFENTNDIJDGETVYISGASCTTNCIAPAKALHDAFGIOKIMTTIAY 181
 DB 121 PANEEDITTVGVNENKRYANAHNDVSNASCTTNCIAPAKALHDAFGIOKIMTTIAY 180
 QY 182 TSDQMIIDGPHRGDULRRARAGANIVNNGAAXIGVITPELNGKIDGAAQRPVPTG 241
 DB 181 TMDQRIQDAEPH-NDLRARAAAANIVITTTGAAKAIGLVEANGLKNGSLKRVPLTG 238

Thu May 6 08:16:12 2004

us-09-878-781-4.rsp

Page 12

[illegible]

Search completed: May 5, 2004, 14:20:11
Job time : 18 secs